PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12Q 1/68 — A3

(11) International Publication Number: **WO 99/34016**

(43) International Publication Date: 8 July 1999 (08.07.99)

(21) International Application Number: PCT/IL98/00625

(22) International Filing Date: 28 December 1998 (28.12.98)

(30) Priority Data:

122793	29 December 1997 (29.12.97)	IL
126627	16 October 1998 (16.10.98)	IL

(71) Applicant *(for all designated States except US)*: GENENA LTD. [IL/IL]; Moshav Cherut, 40600 Tel–Mond Mail (IL).

(72) Inventor; and

(75) Inventor/Applicant *(for US only)*: VIDER, Ben–Zion [IL/IL]; Moshav Cherut, 40600 Tel–Mond Mail (IL).

(74) Agent: REINHOLD COHN AND PARTNERS; P.O. Box 4060, 61040 Tel Aviv (IL).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
With international search report.

(88) Date of publication of the international search report: 2 September 1999 (02.09.99)

(54) Title: A METHOD FOR IDENTIFYING AND CHARACTERIZING CELLS AND TISSUES

(57) Abstract

A method for determining the genetic proximity of two cells. The pattern of expression of genes in a selected family is determined in both cells and is a proximity index calculated based upon the patterns of gene expression. A large proximity index indicates that the two cells are genetically proximal to each other, while a small proximity index indicates that the two cells are genetically distant from each other.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

A METHOD FOR IDENTIFYING AND CHARACTERIZING CELLS AND TISSUES

FIELD OF THE INVENTION

The invention relates to diagnostic methods and more specifically to methods for characterizing cells.

BACKGROUND OF THE INVENTION

There are many situations in which it is important to be able to characterize a cell so as to be able to identify it. For example, during metastasis, tumorigenic cells leave their tissue of origin and migrate to other locations in the body where they are capable of forming secondary tumors. These tumors are thus derived from cells foreign to the surrounding host tissue. Effective treatment of these tumors relies, *inter alia*, on identifying the tissue of origin of the cells forming the tumor.

There are also situations when it is important to know whether the genetic status of a cell, (the combination of genes expressed in the cell) has been altered following a particular treatment, e.g. drugs, irradiation, transfection or prolonged culturing.

There are also cases when it is desirable to know whether a cell carries a genetic defect, for example in prenatal detection of genetic defects.

The few methods presently available for characterizing cells in order to determine their origin or genetic status are laborious and require a highly competent person to carry them out (e.g. evaluating histological stainings). Most of these techniques are only able to characterize cells on the basis of aberrations in chromosome morphology.

The homeobox (HB) containing genes comprise a major group of genes known to play a key role in developmental processes. Their gene products, the homeoproteins, all contain a highly conserved 61-amino acid homeobox domain, which forms a helix-turn-helix DNA-binding site. Sequences flanking the homeodomain possess activating or repressing functions.

HB genes are classified by several criteria including sequence homology within and adjacent to the homeodomain (HD), the developmental stages at which they are expressed, and the tissues in which they are expressed. The *Drosophila* homeobox genes have been shown to have a sequential pattern of expression during embryonic development, i.e. a specific set of homeobox genes are expressed at each developmental stage. The mammalian homologues of the single *Hox* gene cluster in *Drosophila* correspond to four mammalian *Hox* gene clusters which is the largest group of mammalian HB genes. These genes form four clusters termed *Hox* A, B, C and D located on four different chromosomes and are characterized in having a size of less than 5 kB and comprising a single intron. Other groups of mammalian homeobox genes are also dispersed on different chromosomes but are typically larger (more than fifty kB) and have several introns.

The involvement of the mammalian homeobox genes in various processes (e.g. malignancy) has been suggested (Vider *et al.*, *Biochem. Biophys. Research Communications* 232:742-748, 1997). However, to date, the analysis of the expression of genes belonging to this family has typically focused on the expression of a specific gene in various cells or under different conditions.

Enzymes of the eukaryotic protein kinase superfamily catalyze the reversible transfer of the γ-phosphate from ATP to the amino acid side chains of proteins. The state of phosphorylation of a protein can have profound effects on its activity and its ability to interact with other proteins. Protein

kinases are thus involved in many aspects of cellular regulation and metabolism. The catalytic domains of the eukaryotic protein kinases are 250 to 300 amino acids in length with 12 highly conserved subdomains. Phylogenetic analysis of the catalytic domains has revealed five main groups in the protein kinase super family.

The mitogen-activated protein kinase (MAPK) family is one family in the serine threonine kinase super family. This family is involved in the signal transduction pathways of cell cycle regulated events. MAPKs are extracellular signal-regulated kinases (ERKs), that are activated by phosphorylation on their theonine and tyrosine residues. Protein tyrosine kinases (PTK) are also known to play important roles in oncogenesis. Many PTKs form the cytoplasmic moiety of various membrane receptors.

The present invention provides a method for characterizing a cell by means of a genetic proximity index. The method is based upon the pattern of expression of genes in a selected gene family. The method of the invention may be used for example for determining the origin of a cell, its genetic status, whether it carries a genetic defect, or whether it is transformed.

SUMMARY OF THE INVENTION

In the following description and set of claims the term *"gene family"* will denote a set of genes present in the cells of an organism whose gene products have an homology with one another of at least 70%. Examples of such gene families are the homeobox gene family and the kinase gene family.

The term *"pattern of gene expression"* of a selected gene family in a cell will refer to a set of genes of the gene family expressed in the cell.

The present invention is based on the novel finding that a particular cell of an organism expresses a unique combination of genes from a selected gene family. This pattern of gene expression thus serves as a signature of the cell. The expression pattern of the gene family in a particular cell will

sometimes herein be referred to as the genetic status of the cell. Thus, in accordance with the invention, it was shown that cells originating from a specific organ demonstrate a characteristic pattern of expression of genes from a selected gene family while cells originating from organs of different embryonic origin express a different pattern of expression of genes from the same gene family.

The method of the invention is based on comparing the expression pattern of a selected gene family in a first cell to that in a second cell. Generally, the number of genes in the gene family expressed in both of the two cells is compared to the total number of genes in the family expressed in at least one of them. The larger the number of genes from the selected gene family expressed in both cell types, the greater the likelihood that the first and second cells are of a similar origin.

For the first time, in accordance with the present invention, the expression pattern of genes in a selected family of a first cell is compared with that of a second cell by means of a *"proximity index"* which enables a person versed in the art to easily compare various characteristics of the two cells including for example their origin and genetic status.

For the calculation of the proximity index between a first cell and a second cell for a selected gene family, the expression level of each gene in the gene family is quantitatively determined for each of the two cells, for example, as described in the examples below. The proximity index I between the two cells for the selected gene family is then calculated according to the expression:

$$I = \frac{\sum_i \min(a_i, b_i)}{\sum_i \max(a_i, b_i)}$$

wherein a_i and b_i are the expression level of the gene i in the first and second cell, respectively, and the summations are performed over all genes i in the gene family.

The proximity index I is a number between 0 and 1, inclusive. A proximity index of 0 means that the two cells do not express any genes from the selected family in common, indicating for example that the two cells are of different origins. An index of 1 means that the cells express an identical set of genes from the selected family, indicating for example a high probability that the two cells have a common origin.

The present invention thus provides a method for determining the genetic proximity of a first cell and a second cell comprising the steps of:

(a) obtaining said first cell and said second cell;

(b) determining in said first cell and said second cell the pattern of expression of genes in a selected gene family;

(c) calculating a proximity index I, wherein

$$I = \frac{\sum_i \min(a_i, b_i)}{\sum_i \max(a_i, b_i)}$$

wherein a_i and b_i are the expression level of the gene i in the first and second cell, respectively, and the summations are performed over all genes i in the gene family.

A large proximity index indicating that said first cell and said second cell are genetically proximal to each other, a small proximity index indicating that said first cell and said second cell are genetically distant from each other.

The expression pattern of the selected gene family in one of the two cells may be determined before the method of the invention is carried out

on the second cell. In this case, the expression pattern of the first cell may be provided in accordance with the invention e.g. in the form of a catalogue. The expression pattern of the second cell when subsequently determined by the method of the invention may be compared to that of the first cell provided in the catalogue so as to allow calculation of the proximity index of the two cells. It should be noted that, while the expression pattern of the genes in a given family is essentially the same in all cells of a particular origin, there may be some variability, for example in cells obtained from individuals of different genetic backgrounds. Thus, when using a predetermined expression pattern of a cell minor adaptations may occasionally be necessary.

Alternatively, the gene expression pattern of the two cells may be determined simultaneously by the method of the invention.

In order to determine the pattern of expression of genes from the selected gene family in a cell, any of the methods known in the art for detecting gene expression in cells may be used, such as for example one of the methods described in Sambrook *et al.* (Sambrook *et al.* in: Molecular Cloning, A Laboratory Manual, Coldspring Harbour Laboratory, Coldspring Harbour Lab. Press, USA, 1989). In addition, there are a number of commercially available kits which may be used to determine the genetic expression pattern of cells in accordance with the invention.

By a preferred embodiment, gene expression is determined in tested or reference cells using the reverse transcriptase (RT) PCR (RT-PCR) method. RNA samples from the tested cells are first reverse transcribed and the cDNA products are then used as templates for PCR using conserved primers or primers specific for each of the genes.

Typically, a first stage of the method of the invention involves rapid screening for gene expression followed by a second stage involving a more specific analysis. For the rapid screening, PCR is performed in the presence of primers designed to complement conserved regions in the gene

family. In the case of the HB genes, for example, the primers used may be complementary to regions of the genes or to regions of dispersed homeobox genes.

For the specific analysis, PCR is performed using sense primers specific to a given gene in the family and internal to the upstream sense primers used in the rapid screening stage. The conserved antisense primers used in the first screening stage may also be used in the second specific stage if either the sequence upstream or downstream to the gene is specific. A more detailed explanation of the above two stages used in accordance with the invention to determine the genetic status of a tested cell is explained in more detail below in the Examples. Lists of the primers used in accordance with the invention are also provided below. However, the primers included in the lists should not be construed as limiting and any primer as defined above may be used in the method of the invention.

The relative expression level of genes in a cell used in the calculation of the proximity index may be determined for example by comparing the level of expression of the genes to that of a standard gene. Alternatively, the expression level of the genes may be determined as follows. Sense and antisense primers designed to match conserved sequences in the gene family are used to obtain PCR products of several different genes in the gene family. These PCR products are ligated to a cloning vector and propagated in an appropriate bacterial host. The colonies are isolated and the clones sequenced. The total number of clones sequenced in each of the two cells must be the same. The identity of the PCR products is determined and the number of clones of each gene represents its relative expression in the PCR reaction. When the proximity index I between two cells is calculated from relative expression levels obtained by this procedure, a_i and b_i are the number of clones of gene *i* isolated from the first and second cell, respectively. Sequencing the PCR products may optionally be simplified by utilizing the fact

that within conserved sequences there are nonetheless nucleotides specific for each particular gene. It is thus only necessary to sequence a fragment of about 30 bp within the sequence to be able to identify every PCR product by using databank sequences. A sequencing procedure capable of separating 200-300 bp is thus able to analyze about 10 different stranded PCR products.

In accordance with the invention it has also been found that the genetic status of a cell as determined by the expression pattern of a selected gene family may undergo specific and detectable changes following exposure of the cells to a given treatment such as irradiation, drugs, exposure to growth factors, gene transfection, etc. The genetic status of a cell may also change spontaneously, for example, during the transition from the fetal stage to the adult stage. The method of the invention may therefore be used for determining whether a treatment has altered the genetic status of a cell by calculating the proximity index of the cell before having been exposed to the treatment and the same cell after having been exposed to the treatment.

The present invention is based on the further finding that the set of genes in a given family expressed by a transformed cell may be different from that of its untransformed counterpart. The invention may therefore be used for determining whether a cell is transformed by calculating the proximity index of a cell and its untransformed counterpart.

In accordance with the invention, it has also been found that the expression pattern of a selected gene family in a cell may be abnormal in cases where the cell carries a genetic defect in its genome. The method of the invention may therefore be used for detecting the presence of a selected genetic defect in the genome of a cell by calculating the proximity index of the cell and a cell of the same type not having the genetic defect. In one embodiment of the invention, a genetic defect in an unborn fetus is detected by calculating the proximity index of an amniotic cell of the fetus obtained for example by

amniocentesis and a cell of the same type obtained from a fetus not having a genetic defect.

By yet another aspect of the present invention, it has been realized that cells expressing a particular set of genes in a given gene family may possess a specific phenotype. Thus, it may be possible to alter the phenotype of a cell by transfecting it with a set of genes known to give rise to a desired phenotype. For example, it is known that placental cells are capable of producing and secreting a number of hormones useful in various therapeutic applications. Following identification of the set of genes from a selected gene family expressed in these cells, it is possible, in accordance with the invention, to transfect with the identified set of genes cells that normally do not produce these hormones or produce them in small quantities so as to obtain cells capable of producing the desired hormones in large quantities. In other cases, the transfection of a specific set of genes into target cells may result in their transdifferentiation into cells having a desired property. For example, human placental cells may be transfected with a set of HB genes causing them to differentiate into adrenal or kidney cells having close genetic proximity. This would be useful, for example, for maintaining tissues in a tissue or organ bank. The tissue may be maintained in the form which is easiest to grow or maintain in culture and made to transdifferentiate into the desired tissue according to transplantation needs.

Thus, by yet another aspect of the invention, there is provided a method for obtaining cells capable of expressing an HB related desired property comprising the steps of:

(a) identifying a specific pattern of expression of HB genes in cells having a desired property;

(b) transfecting said identified set of HB genes into target cells lacking said desired property under conditions enabling expression of said HB

cells in said target cells, said transfection resulting in expression of said desired property in said transfected cells.

By another of its aspects, the present invention provides a kit for carrying out the method of the invention comprising:

(a) means for obtaining the mRNA from said first cell and said second cell;

(b) means for performing the reverse transcriptase polymer chain reaction on the mRNA obtained from said first cell and said second cell;

(c) means for detecting the genes of said selected gene family expressed in said first cell and in said second cell;

(d) instructions for carrying out the method.

The invention also provides a kit for carrying out the invention comprising:

(a) means for obtaining the mRNA from said first cell;

(b) means for performing the reverse transcriptase polymer chain reaction on the mRNA obtained from said first cell;

(c) means for detecting the genes of said selected gene family expressed in said first cell;

(d) a catalogue providing the genes of said selected gene family expressed in said second cell;

(e) instructions for carrying out the method.

BRIEF DESCRIPTION OF THE FIGURES

The invention will now be demonstrated by way of non-limiting examples with reference to the following figures in which:

Figs. 1 A-D are photographs showing the expression pattern of the conserved HOX PCR products detected using the primers shown in List 1 below, in which (A) is from placenta, (B) is from *S. nigra*, (C) is from normal colon and (D) is from tumorous colon. The primers used were in Lane 1 HOX

1(+) and HOX 1(-), in Lane 2 HOX 1(+) and HOX 2(-), in Lane 3 HOX 2(+) and HOX 1(-), in Lane 4 HOX 2(+) and HOX 2(-), in Lane 5 HOX 3(+) and HOX 1(-), in Lane 6 HOX 3(+) and HOX 2(-), and M is the molecular weight standards.

Figs. 1 E-G are photographs showing the expression pattern of the conserved HOX PCR products detected using the primers shown in List 4. The OCT 1* and PBX* PCR products also separated on these gels. E- Adrenal, F- Fetal Brain, G- Adult Brain. The primers used were, Lane 1: 1(+)1(-), Lane 2: 1(+)2(-), Lane 3: 2(+)1(-), Lane 4: 2(+)2(-), Lane 5: 3(+)1(-), Lane 6: 3(+)2(-), Lane 7: OCT 1, Lane 8: PBX 1, Lane 9: marker Low DNA Mass Ladder, Lane 10: ϕX174RF DNA/Hae III.

Figs. 2 A-C are photographs showing the expression pattern of several HOX genes obtained by reverse transcriptase of RNAs using specific HOX oligo –dT primers (List 2) in which A, B, C and D are as described above for Fig. 1. The primers used were in Lane 1 is A5, Lane 2 A7, Lane 3 A11, Lane 4 A13, Lane 5 B1, Lane 6 B2, Lane 7 B3, Lane 8 B6,7, Lane 9 C5, Lane 10 C8, Lane 11 C12, Lane 12 C13, and Lane 13 D3.

Figs. 3 A-B are photographs showing the expression pattern of dispersed homeobox genes using primers of List 3 in which A and B are as described above for Fig. 1. The primers used Lane 1 is EVE2, Lane 2 MSX1, Lane 3 MSX2, Lane 4 OCT, Lane 5 PAX, Lane 6 CFB1, Lane 7 LFB3, Lane 8 Bicoid, Lane 9 Goosecoid, Lane 10 Engrailed, Lane 11 EMX, Lane 12 OTX, Lane 14 OTX, Lane 14 CDX, Lane M PUC 19 Hea II.

Figs. 3C-G are photographs showing the expression pattern of several other dispersed homebox genes using primers of list 3.

C- Placenta: Lane 1 is OCT 1*, Lane 2 is PBX 1*, Lanes 3 and 4 are empty, Lane 5 is Bicoid, Lane 6 s Engrailed, Lane 7 is DLX a, Lane 8 is DLX b, Lane 9 is DLX c, Lane 10 is DLX d, Lane 11 is GBX, Lane 12 is MEI,

Lane 13 is P HOX, Lane 14 is PROX 1, Lane 15 is marker Low DNA Mass Ladder. Land 16 is ϕX174 RF DNA/Hae III.

D- Fetal Kidney: Lane 1 is OCT 1*, Lane 2 is PBX 1*, Lane 3 is OCT, Lane 4 is CDX, Lane 5 is Bicoid, Lane 6 is Engrailed. Lane 7 is DLX a, Lane 8 is DLX b, Lane 9 is DLX c, Lane 10 is DLX d, Lane 11 is GBX, Lane 12 is MEI, Lane 13 is P HOX, Lane 14 is PROX 1, Lane 15 is Marker Low DNA Mass Ladder, Lane 16 is ϕX174 RF DNA/Hae III.

E- Adrenal Gland: Lane 1 is OCT, Lane 2 is BRN a, Lane 3 is BRN b, Lane 4 is CDX, Lane 5 is Bicoid, Lane 6 is Engrailed, Lane 7 is DLX a, Lane 8 is DLX b, Lane 9 is DLX c, Lane 10 is DLX d, Lane 11 is GBX, Lane 12 is MEI, Lane 13 is P HOX, Lane 14 is PROX 1, Lane 15 is marker Low DNA Mass Ladder, Lane 16 is ϕX174 RF DNA/Hae III.

F- Adult Brain: Lane 1 is OCT, Lane 2 is BRN a, Lane 3 is BRN b, Lane 4 is CDX, Lane 5 is Bicoid, Lane 6 is Engrailed, Lane 7 is DLX a, Lane 8 is DLX b, Lane 9 is DLX c, Lane 10 is DLX d, Lane 11 is GBX, Lane 12 is MEI, Lane 13 is P HOX, Lane 14 is PROX 1, Lane 15 is marker Low DNA Mass Ladder, Lane 16 is ϕX174 RF DNA/Hae III.

G- *S. Nigra*: Lane 1 is OCT 1, Lane 2 is PBX, Lane 3 is OCT, Lane 4 is BRN a, Lane 5 is BRN b, Lane 6 is CDX, Lane 7 is Engrailed, Lane 8 is DLX a, Lane 9 is DLX b, Lane 10 is DLX c, Lane 11 is DLX d, Lane 12 is GBX, Lane 13 is MEI, Lane 14 is P HOX, Lane 15 is PROX 1, Lane 16 is marker Low DNA Mass Ladder.

Figs. 4 A-F are photographs showing the expressin pattern of the conserved MAP kinase PCR products detected using the primers shown in List 1, in which (A) is Cytogenetic normal amniotic cell culture, (B) is Normal Colon, (C) is tumorigenic Colon, (D) is Adrenal, (E) is Fetal Brain, and (F) is Adult Brain. The primers used were in Lane 1 a(+) and a(-), in Lane 2a(+) and b(-), in Lane 3 a(+) and c(-), in Lane 4 b(+) and a(-), in Lane 5 b(+) and b(-), in Lane 6 b(+) and c(-), in Lane 7 c(+) and a(-), in Lane 8 c(+) and b(-), in Lane

9c(+) and c(-), in Lane 10 d(+) and a(-), in Lane 11 d(+) and b(-), in Lane 12 d(+) and c(-), Lane 13 is marker Low DNA Mass Ladder. Lane 14 is ϕX174 RF DNA/Hae III.

Figs. 5 A-G are photographs showing the expression of the pattern of the conserved PTK kinase PCR products detected using the primers shown in List 1, in which (A) is Placenta, (B) Cytogenetic normal amniotic cell culture (C) is Normal Colon, (D) is Tumorigenic Colon, (E) is Adrenal, (F) is Fetal Brain and (G) is Adult Brain. The primers used were in Lane 1 e(+) and e(-), in Lane 2 e(+) and f(-), in Lane 3 e(+) and g(-), in Lane 4 e(+) and h(-), in Lane 5 f(+) and e(-), in Lane 6 f(+) and f(-), in Lane 7 f(+) and g(-), in Lane 8 f(+) and h(-), in Lane 9 g(+) and e(-), in Lane 10 g(+) and f(-), in Lane 11 g(+) and g(-), in Land 12 g(+) and h(-), in Lane 13 h(+) and e(-), in Lane 14 h(+) and f(-), in Lane 15 h(+) and g(-), in Lane 16 is marker Low DNA Mass Ladder.

Fig. 6 is a photograph showing the expression pattern of several dispersed homeobox genes using primers of List 3 in amniotic cell cultures. Lane 1 is EVE 1, Lane 2 is EVE 2, Lane 3 is MSX 1, Lane 4 is MSX 2, Lane 5 is OCT, Lane 6 is OCT 1, Lane 7 is BRNa, Lane 8 is BRNb, Lane 9 is LFB3, Lane 10 is LFB1, Lane 11 is PAX, Lane 12 is Bicoid, Lane 13 is Goosecoid, Lane 14 is Engrailed, Lane 15 is CDX, Lane 16 is marker Low DNA Mass Ladder.

Figs. 7 A-F are photographs showing the expression pattern of the conserved MAP kinase PCR products detected using the primers shown in List 4, in amniotic cell cultures: in which (A) is culture 1, (B) is culture 3, (C) is culture 4, (D) is culture 5, (E) is culture 6, (F) is culture 7. The primers used were in Lane 1 a(+) and a(-), in Lane 2 a(+) and b(-), in Lane 3 a(+) and c(-), in Lane 4 b(+) and a(-) and a(-), in Lane 5 b(+) and b(-), in Lane 6 b(+) and c(-), in Lane 7 c(+) and a(-), in Lane 8 c(+) and b(-), in Lane 9 c(+) and c(-), in Lane 10 d(+) and a(-), in Lane 11 d(+) and b(-), in Lane 12 d(+) and c(-), Lane 13 is marker Low DNA Mass Ladder. In addition, lanes 14-16

contained in several photographs the following PCR products: Lane 14 is OCT1, Lane 15 is PBX, Lane 16 is β Actin.

Figs. 8 A-G are photographs showing the expression pattern of the conserved PTK kinase PCR products detected using the primers shown in List 5, in amniotic cell cultures, in which (A) is culture 1, (B) is culture 2, (C) is culture 3, (D) is culture 4, (E) is culture 5, (F) is culture 6, (G) is culture 7, (H) is culture 8. The primers used were in Lane 1 e(+) and e(-), in Lane 2 e(+) and f(-), in Lane 3 (+) and g(-), in Lane 4 e(+) and h(-), in Lane 5 f (+) and e(-), in Lane 6 (+) and f(-), in Lane 7 f(+) and g(-), in Lane 8 f(+) and h(-), in Lane 9 g(+) and e(-), in Lane 10 g(+) and f(-), in Lane 11 g(+) and g(-), in Lane 12 g(+) and h(-), in Lane 13 h(+) and e(-), in Lane 14 h(+) and f(-), in Lane 15 h(+) and g(-), in Lane 16 is marker Low DNA Mass Ladder.

Figs. 9 A-C are photographs showing the expression pattern of the conserved Serine Threonine kinase PCR products, detected using the primers designated STK 8-21 (List 6), in amniotic cell cultures: in which (A) is culture 1, (B) is culture 3, (C) is culture 4. The primers used were in Lane 1 i(+) and j(-), in Lane 2 j(+) and j(-), in Lane 3 k(+) and j(-), in Lane 4 l(+) and j(-), in Lane 5 m(+) and j(-), in Lane 6 n(+) and j(-), in Lane 7 o(+) and j(-), in Lane 8 i(+) and k(-), in Lane 9 j(+) and k(-), in Lane 10 k(+) and k(-), in Lane 11 l(+) and k(-), in Lane 12 m(+) and k(-), Lane 13 n(+) and k (-), Lane 14 o(+) and k(-), Lane 15 is marker Low DNA Mass Ladder.

EXAMPLES

Materials and Methods

Human poly A RNAs were obtained from Clontech. Reverse transcriptase (RT) – Polymerase Chain Reaction (PCR) was performed as follows. 1 μg samples of poly A RNA or total RNA were reverse-transcribed using recombinant AMV reverse transcriptase in the presence of oligo-dT as an anti-sense primer. 3μl of the reverse-transcribed mixture was used as a template for PCR, using AmpliTaq polymerase in the presence of the conserved primers

shown in List 1 or the specific primers shown in Lists 2 and 3. Amplification was performed in an Eppendorf Master Cycler personal PCR instrument for 35 cycles. Each cycle consisted of 1 min. of denaturation at 94°C, followed by 1 min. of annealing usually at 55°C, and 1 min. of extension at 72°C. The same cDNAs were used as templates for parallel PCR reactions performed for 20 cycles in the presence of the β-actin primers:

5'-GTTGAGACCTTCAACACCCC-3'

5'-GTGGCCATCTCTTGCTCGAAGTC

RNA samples from various cell populations were reverse transcribed. The cDNA produced from each sample, was used as a template for PCR using conserved and specific primers for each of the genes studied. Since genes belonging to the same group have specific domains in common such as the *paired pou* and *lim* domains in HB genes, it was possible to use a single primer in each PCR reaction, instead of two. The RT-PCR products were analyzed by electrophoresis on 2% or 3% high resolution gels (Ambion), and photographed.

Amniotic cells cultures were provided by the Cytogenetic Unit of the Herzlia Medical Center. 5 cytogenetically normal cultures were used (cultures 1-3, 5 and 6 together with cultures of trisomy 21 cells (cultures 4 and 8) and a culture of cells showing balanced translocation 1:13 (culture 7) from a phenotypically normal fetus. The cultures are described in Table 10. After 3 weeks of incubation, RNA was extracted using the Total RNA Isolation Kit (Ambion). Cultures were analyzed by RT-PCR utilizing the primers given in Lists 1-10.

EXAMPLE 1

The following tissues were analyzed by the method of the invention:

1. Human placenta – the placenta is derived primarily from embryonic trophoblast and mesodermal cells. This organ is considered to have embroyogenic proximity to epithelia derived from ectoderm, neural derived organs such as the adrenal medulla as well as mesodermal derived tissues such as the heart, hemangioblastic tissue, adrenal and cortex.
2. Human substantia nigra – This part of the midbrain is derived from neuroectoderm.
3. Normal and malignant colonic tissues – As part of the hindgut, the colon is derived from endoderm.

The study was divided into two stages, the first stage consisted of a rapid screening method, designed to determine whether a particular homeobox group is expressed in the examined organ. The second stage consisted of a secondary PCR in order to determine the specific genes expressed in the given homeobox group.

An oligo d(T) primer was used to synthesize cDNA from the polyA RNA. PCR reactions were then performed in the presence of primers designed to complement conserved regions of the homeobox groups (List 1). The degree of similarity between the primers and the conserved regions was 85% or more, with no mismatch in the last 2 nucleotides at the 3′ termini between the primer and the matched sequence. List 2 shows the pairs of primers used and some of the HOX genes transcribed by them. The PCR products were analyzed by gel electrophoresis, stained with ethidium bromide and photographed under UV light.

Results

The results of the screening stage are shown in Fig. 1 A-D. *S. Nigra* expresses low levels of HOX genes, with the genes at the 5′ end in all clusters being not expressed at all. Human placenta and both normal and malignant colonic tissues express high levels of the HOX genes. Comparing the relative intensities of the bands shows that the placenta and colonic tissues possess different expression profiles. Normal and malignant colonic tissues have similar expression profiles although some differences are evident. For example, the pair of primers HOX 2(+), HOX 2(-) (Lanes 4, C and D in the Fig.). Each of the 4 samples originating from 4 organs studied thus have a unique expression profile with the pairs of primers used.

PCR was then performed using specific sense primers upstream to the homeodomain but internal to the upstream sense primers used in the screening stage (List 2). For each gene, a specific sense primer was used together with the most matched anti-sense primer from the 3 primers used in the screening stage. In order to ensure that different nucleotides appear at the 3′ end of the specific primers, and that the primer matches only the desired HOX gene, specific HOX gene PCR was performed on cDNAs taking advantage of the conserved anti-sense primers, and the ability to compare results between the screening stage and specific stage in each organ.

The PCR products were then separated by gel as shown in Fig. 2. The expression level of each HOX gene was determined for each tissue and the results are summarized in Table 1. Comparison of Figs. 1 and 2 shows that there is a good correlation between the results obtained in the screening stage (Fig. 1) and those obtained with the specific HOX gene primers in each organ (Fig. 2). Moreoever, samples obtained from various organs tested demonstrate a characteristic pattern of expression of individual HOX genes. For example, organs having completely different embryonic region Nigra-(ectoderm),

Colon-(endorderm), Placenta-(mesoectoderm) have a different HOX gene expression pattern.

The proximity of two tissues is given by the homeobox index I which is defined and calculated as explained above.

Several examples of the proximity index obtained by the above method on the data of Table 1 are as follows:

Placenta – *S. Nigra* = 4.5/10 = 0.45

Placenta – Colon normal = 4/11.5 = 0.34

S. Nigra – Colon normal: 3.5/9.5 = 0.36

Colon Normal – Colon tumor = 7/8.5 = 0.82.

EXAMPLE 2

PCR was also performed on the samples used in Example 1 above using primers to the dispersed HB genes (List 3). These primers contained either both sense and antisense from conserved regions of HB groups (PAX, OCT, CDX), or contained one specific primer (EVE, MSX etc.). The PCR products were separated by gel and some of the products are shown in Fig. 3 A-B (from Placenta and *S. Nigra*).

As can be seen in Table 2, the expression pattern of these genes was also unique for each different tissue.

EXAMPLE 3

The proximity between cells originating from placenta, adrenal and fetal kidney was calculated on the basis of the expression level of HOX genes (Table 3a) and dispersed homeobox genes (Table 3b) as described in Example 1 above.

On the basis of the results shown in Table 3a and b, the proximity of the cells of the three above mentioned organs was calculated as explained above to give the following results:

Fetal kidney and adrenal: 7.5/12 = 0.62

Placenta and fetal kidney: 5.5/12 = 0.46

Placenta and adrenal: 5.5/10.5 = 0.52

These results demonstrate a higher degree of proximity between the adrenal gland and F. kidney than between placenta, *S. Nigra*, and colon. It thus reflects the higher degree of embryogenic proximity of these organs including their final position in the body (Kidney and Adrenal) than that of *S. Nigra*, placenta and colon.

EXAMPLE 4

The following tissues were analyzed by the method of the invention.

1. Adult kidney (A. kidney)
2. Fetal brain (F. brain)
3. Adult brain (A.brain).

The HB genes screened are as follows:

a. I. HOX conserved primers II. HOX specific primers

b. POU domain containing genes:
I. OCT conserved primers II.OCT primers III. LFB specific primers
IV. BRN specific primers

c. *Drosophila* like maternally expressed genes I. CDX conserved primers. II. CDX specific primers. III Bicoid specific primers.

d. *Drosophila* like segmented expressed genes. I. PAX conserved primers. II. PAX specific primers. III Engrailed specific primers. IV Goosecoid specific primers.

e. I. LIM domain containing genes conserved primers. LIM domain containing specific primers.

f. I. NK genes conserved primers. II. NK genes specific primers.

g. DLX genes conserved primers.

f. CNS embryonic expressed genes: EMX, OTX.

h. Miscellaneous HB genes: EVE1,2 MSX1,2 GBX, MEI 1, PHOX, PROX 1, PBX 1. The primers sequences are given in Lists 4, 5 and 6.

Three levels of gene expression were distinguished instead of just two as used in Example 1. The level of expression was referred to the levels of the homeobox genes PBX and OCT-1 (very low, regular, and high in comparison to the expression levels of OCT 1, and PBX1). A very low expression level was a level barely seen on the gel, and it is socred as 1. The regular level is any expression less than or equal to that of PBX or OCT 1 and is scored as 2. Levels greater than or equal to that of PBX 1 or OCT 1 (the maximum of the two) is scored as 3. PBX 1 and OCT 1 were found by the methods of the invention to be expressed in all cells examined. They have been suggested to work in concert with HOX genes, in order to tune and achieve their final effect.

Results

The expression pattern and scoring of PCR products is given in Tables 4-6.

The verification of the desired product was made by electrophoresing the PCR product on the gel, and deforming its molecular weight using molecular weight standards. In several cases secondary PCR was made on the primary PCR with internal primers, in order to confirm the molecular weight. PCR with internal primers was also used in instances where the primary PCR primers were designed to match conserved regions transcribing many genes that belong to the same group. Then a secondary PCR was performed with one specific internal primer that could match only one gene of that group.

Results of the scoring of the PCR products are summarized in Figs. 1 E-G and 3 C-G and Tables 4-5.

All organs expresses HOX 7 and B5 and either B5 or B7. This might imply a central role of these genes in the cells. Brain samples express very low levels of HOX genes and a limited number of other genes. On the other hand, a similar pattern of expression using the conserved primers was observed in these brain samples.

HOX genes are more variably expressed in various physiological or pathological states of an organ relative to the dispersed homeobox genes.

There is a high level of expression of CDX in *S. nigra* with the conserved primers and of CDX-4 with the specific primer. A wide distribution of LFB3 expression, expression of ENGRAILED in brain and adrenal gland, and expression of PAX in adult kidney are also observed.

Comparison of the I score demonstrated the following ranges of scoring: low scoring of 0.24 between tumorous colon and adult brain, and the highest scoring 0.83 between normal and tumorous colonic tissues. Adrenal and fetal kidney demonstrated I score of 0.68 consistent with the known biological proximity of these organs. Placenta demonstrated a uniform score of 0.4-0.5 I score with most of the other cells except for brain in which it was lower scoring.

EXAMPLE 5

The method of the invention was performed using primers from the 11 conserved subdomains of the kinase family catalytic domain. The PCR products were separated by gel electrophoresis. Analysis of the PCR products showed variable expression of the different genes. As above, the extent of expression of a given gene was rated as 1 (very low), 2 (medium) or 3 (high).

The results are given in Figs. 4 and 5 and Tables 7 and 8.

The proximity of two cells is given by the proximity index for kinase index I which is defined and calculated as explained before. The obtained values of the kinase proximity index are summarized in Table 9.

It is clear from Tables 7 and 8 that there is much more variability of the expression of the PTK genes among various organs in comparison to that of the MAPK genes.

EXAMPLE 6

The expression pattern of the homeobox genes in amniotic cell cultures was analyzed in 2 amniotic cell cultures. The CR products of culture 2 separated on gel are shown in Fig. 9. The expression of the homeeobox genes is summarized in Table 11. The BRN and PCR gene products were both expressed (similar to brain), but not that of LFB1 and LFB3. The very low level of PROX expression observed is unusual. On the other hand, the HOX genes expressed at a level greater than that in placenta, for example. Although the placenta and the amniotic sac are both extra embryonic tissues their homeobox proximity index was found to be only 0.49.

EXAMPLE 7

The 8 amniotic cell cultures described in Table 10 were examined for kinase gene expression. The primers given in Lists 7-10 were used to obtain the RT-PCR products of these genes. Figs. 7, 8 and 9, demonstrate partial MAPK, partial PTK, and partial STK expression in these cultures.

Culture 3 does not present any cytogenetic abnormality, however, a PCR product of the MAP 3 PCR product was observed (Fig.7B, lane 3) that was absent from all other cultures (Fig. 7). Culture 3 also demonstrated a unique pattern of PTK expression (Fig. 8C, lanes 9-11) and a very low expression of the STK 11 gene (Fig. 9B, lane 4) in comparison to the other cultures (e.g. Fig. 9A and C, lane 4). Cultures 4 and 8 (trisomy 21) were also the only cultures in which a similar expression pattern of PCR products was obtained using primers PTK 9 and 11 (Fig. 8D and H). This is in contrast to the other cultures in which more PCR products were obtained using PTK 9.

The karyotype of Culture 7 shows a balanced translocation between chromosomes 1 and 13. This translocation is also present in the healthy parents, so this translocation does not lead to a diseased state in the adult. Similarly, in the fetal cells of Culture 7, this translocation did not result in any changes in the expression of MAPK, PTK, or STK genes (Figs. 7F and 8G). Since genetic defects in one organ are often reflected in other organs (a phenomenon known as association in which apparently unrelated genetic anomalies occur together more frequently than would be expected by chance alone) this embodiment can be used to detect future developmental abnormalities. Examples of such associations include abnormal external ears associated with renal anomalies, a single umbilical artery associated with cardiac defects, and wide spaced eyes or no occipital hair were associated with brain malformations. The results obtained on the amniotic fetal cells also show that cells having a normal karyotype may have an aberrant pattern of gene expression. This embodiment of the invention can thus be used to perform a *"genetic physical examination"* on a fetus.

Example 8

The following cells were analyzed by the method of the invention for expression levels of serine-threonine phosphatase, tyrosine phosphatase and POU homeobox genes:

1. Normal colon cells
2. HT-29 cells
3. Caco-2 cells
4. 3T3 cells.

Expression levels were obtained using the primers given in Lists 11 to 14 and 19. Tables 12 to 14 show the expression levels of the genes studied in the 4 cell types used. The data in Tables 12-14 were then used to obtain proximity indexes between pairs of the cell types from the above list.

The data in Table 12 produced the following proximity indexes:

Cell pair	Proximity index
Normal Colon – HT29	**0.67**
Normal Colon – Caco2	**0.78**
Normal Colon – 3T3	**0.57**
HT29 – Caco2	**0.76**
HT29 – 3T3	**0.58**
CaCo2 – 3T3	**0.66**

When the data from Tables 13 and 14 were used to obtain proximity indexes the following results were obtained:

Cell pair	Proximity index
Normal Colon – HT29	**0.73**
Normal Colon – Caco2	**0.79**
HT29 – Caco2	**0.77**

Cell pair	Proximity index
Normal Colon – HT29	**0.65**
Normal Colon – Caco2	**0.72**
HT29 – Caco2	**0.80**

List 1 – Conserved oligonucleotide primers of the HOX genes used for RT-PCR reactions

Designation	Sequence 5'-3'	Position	Product size (bp)	Matching Hox genes
HOX 1 (+)	TACACTCGCTACCAGACCCTGGAG	21-45	135	A6,A7,B6,
HOX 1 (-)	CCGGTTCTGGAACCAGATCTT	156-136		B7,C5
HOX 1 (+)	TACACTCGCTACCAGACCCTGGAG	21-45	129	A9,A13,B8
HOX 2 (-)	CTGAAACCAGATTTTGACCTG	150-130		D9
HOX 2 (+)	AAGCGCTGCCCCTACACCAA	10-30	146	A9,A10,
HOX 1 (-)	CCGGTTCTGGAACCAGATCTT	156-136		A11,C9,D9
HOX 2 (+)	AAGCGCTGCCCCTACACCAA	10-30	140	A9,A13,D9
HOX 2 (-)	CTGAAACCAGATTTTGACCTG	150-130		D10
HOX 3 (+)	GCCCGGACCACCTACACGCG	10-30	146	A4,A6,A7
HOX 1 (-)	CCGGTTCTGGAACCAGATCTT	156-136		B4
HOX 3 (+)	GCCCGGACCACCTACACGCG	10-30	140	A4
HOX 2 (-)	CTGAAACCAGATTTTGACCTG	150-130		
HOX 4 (+)	CTGGAGCTGGAAAAGGAATT	40-59	116	
HOX 1 (-)	CCGGTTCTGGAACCAGATCTT	156-136		
HOX 4 (+)	CTGGAGCTGGAAAAGGAATT	40-59	110	
HOX 2 (-)	CTGAAACCAGATTTTGACCTG	150-130		
HOX 4 (+)	CTGGAGCTGGAAAAGGAATT	40-59	121	
HOX 3 (-)	CAT CCT GCC GTT CTG AAA CCA	161-141		

List 2 – Specific oligonucleotide primers of HOX genes used for RT-PCR reactions

Designation	Sequence 5'-3'	Position	Product size (bp)	Matching Hox genes
A4 (+)	TTCAATCGCTACCTGACCCGC	64-84		A4, D10
A5 (+)	CGTTACCTGACCCGCAGAA	70-88	91	A5
HOX 3 (-)	CATCCTGCCGTTCTGAAACCA	161-141		
A6 (+)	CACTTCAACCGCTACCTGACA	61-80		A6, B4, C8
A7 (+)	CACTTCAACCGCTACCTGACC	61-80	95	A7
HOX 1 (-)	CCGGTTCTGGAACCAGATCTT	156-136		
A9 (+)	TTCAACATGTACCTCACCAGG	64-84		A9, D9
A10 (+)	TTCAATATGTACCTTACTCGA	64-84		A10
A11 (+)	GTCTACATTAACAAAGAGAAG	70-90	92	A11
HOX 3 (-)	CATCCTGCCGTTCTGAAACCA	161-141		
A13 (+)	CGGGAATACGCCACGAATAAA	51-70	100	A13
HOX 2 (-)	CTGAAACCAGATTTTGACCTG	150-130		
B1 (+)	CATTTCAACAAGTACCTGAG	61-80	96	B1
HOX 1 (-)	CCGGTTCTGGAACCAGATCTT	156-136		
B2(+)	AATAAGTACCTGTGCCGGCCA	67-78	84	B2
HOX 2(-)	CTGAAACCAGATTTTGACCTG	150-130		

List 2 (continued)

B3 (+)	AACCGCTACCTGTGCCGGCCT	67-87	90	B3
HOX 1 (-)	CCGGTTCTGGAACCAGATCTT	156-136		
B5 (+)	AACCGCTACCTGACCCGGCGA	67-87		B5
B6 (+)	CACTACAATCGCTACCTGACG	61-80	96	B6, B7
HOX 1 (-)	CCGGTTCTGGAACCAGATCTT	156-136		
B8 (+)	TTTAATCCCTATCTGACTCGT	63-83		B8
B9 (+)	TACCTCACCAGGGACCGTAGGC	74-94		B9
C4 (+)	CGCTACCTGACCCGAAGGAGA	70-90		C4
C5 (+)	CACTTTAACCGCTACCTCACT	61-81	101	C5
HOX 3 (-)	CATCCTGCCGTTCTGAAACCA	161-141		
C6 (+)	TTTCACTTCAATCGCTACCTA	58-78		C6
C8 (+)	AATCCTTATTTGACACGAAAA	66-86	91	C8
HOX 1 (-)	CCGGTTCTGGAACCAGATCTT	156-136		
C9 (+)	GAGTTTCTCTTCAATATGTATTTA	57-77		C9
C12 (+)	CTGGAGGGCGAGTTTCTGG	46-64	105	C12
HOX 2 (-)	CTGAAACCAGATTTTGACCTG	150-130		
C13 (+)	GAGCTAGAGAAGGAATACGCG	43-63	119	C13
HOX 3 (-)	CATCCTGCCGTTCTGAAACCA	161-141		

List 2 (continued)

D3 (+)	AACCGCTACTTGTGCCGGCCG	67-87	90	D3
HOX 1 (-)	CCGGTTCTGGAACCAGATCTT	156-136		
D4 (+)	CATTTTAACAGGTATCTGACAA	61-82	96	D4
HOX 1 (-)	CCGGTTCTGGAACCAGATCTT	156-136		
D8 (+)	AACCCCTATCTGACCAGGAAA	67-87		D8
D10 (+)	CTCACCCGCGAGCGCCGCCTA	76-96	86	D10
HOX 3 (-)	CATCCTGCCGTTCTGAAACCA	161-141		
D11 (+)	TTTTTCTTTAACGTGTACATA	58-78	104	D11
HOX 3 (-)	CATCCTGCCGTTCTGAAACCA	161-141		

List 3 – Oligonucleotide primers from conserved and specific regions of various dispersed homebox gene groups, for RT-PCR reactions

Designation	Sequence 5'-3'	Position	Product size (bp)	Matching homeobox genes
EVE 1.2 (+)	AACCTGCCCGAAACCACCATC		170	EVE 1
EVE 1 (-)	CAGGGGCAGGTGCGATGG			
EVE 1.2 (+)	AACCTGCCCGAAACCACCATC		147	EVE 2
EVE 2 (-)	GGGGTAGGGCAGGCTTCCGGT			
MSX 1 (+)	CCGCTGGGCCATTTCTCG	760-780	350	
MSX 1.2 (-)	CGGGCTGCGGTTCTGGAACCA	2190-2170		
MSX 2 (+)	AATTCAGAAGATGGAGCGGCG	548-568		
MSX 1,2 (-)	CGGGCTGCGGTTCTGGAACCA	2190-2170		
PAX (+)	GGCTGTGTCAGCAAAATTCT			
PAX (-)	CAGCCTGTCTCGGATCTCCCA			
Bicoid (+)	ATGTCGTCCAGCATGGTGCCC	1110-1130	172	
Bicoid (-)	CGAGTTACACGTGTCCCTATA	1280-1260		
Goosecoid(+)	ACCATCTTTACGGAGGAGCA	1330-1350	143	
Goosecoid(-)	GTTCTTGAACCACACCTC	1473-1455		
Engr 1,2 (+)	AAGATTTGGTTCCAGAACAA	180-200	96	
Engr 1,2 (-)	GTTGTACAGGCCCTGTGCCAT	275-255		
EMX 1.2 (+)	AAACGCATCCGGACCGCCTT	1460-1479		
EMX 1,2 (-)	TCGGTTCTGGAACAACACCTT	1610-1590		
OTX 1,2 (+)	TATCCGGACATATTCATG	1155-1173		
OTX 1,2 (-)	CTGTTGGCGGCACTTGGC			

List 3 (continued)

OTX 1,2 (+)	TATCCGGACATATTCATG	1155-1173		
OTX 1,2 (-)	TTTGCGCTTCTTCCATTT			
CDX 1-4 (+)	GAGCTGGAAAAGGAGTTTCA			
CDX 1-4 (-)	TCCTTGGCTCTGCGGTTCTG			
OCT (+)	CTGAGCTTTAAGAACATGTG	1000-1020	290	OCT-1, OCT-2
OCT (-)	CTGGCGCCGGTTACAGAACCA	1290-1270		
OCT-1 (+)	AGCCCAAGTGCCCTGAATTCT	1077-1097		OCT-1
OCT (-)	CTGGCGCCGGTTACAGAACCA	1290-1270		
OCT-2 (+)	AGCCCCAGCCTGGGTTTC	1021-1038		OCT-2
OCT (-)	CTGGCGCCGGTTACAGAACCA	1290-1270		
BRN (+)	ACCCTGTATGGCAACGTGTT			
BRN (-)	CCCCTTGAGGCTCACCTCGAT			
LFB1 (+)	TCCGGCGACGAGGGCTCCGAG	420-440	663	
LFB (-)	GCGGTTGGCAAACCAGTTGTA	1102-1082		
LFB3 (+)	TCGGAAGATGACACGGATGAC	210-230	790	
LFB (-)	GCGGTTGGCAAACCAGTTGTA	1102-1082		

List 4 - a. I. HOX conserved primers

Conserved oligonucleotide primers of the HOX genes for RT-PCR reactions

Designation	Sequence 5'-3'	Position [b]	Product size (bp)	Matching the HOX genes:	Ref
HOX 1 1(+)	TACACTCGCTACCAGACCCTGGAG	21-45	135	A6,A7,B6	1
1(-)	CCGGTTCTGGAACCAGATCTT	156-136		B7,C5	2
HOX 2 1(+)	TACACTCGCTACCAGACCCTGGAG	21-45	129	A9,A13,B8	1
2(-)	CTGAAACCAGATTTTGACCTG	150-130		D9	2
HOX 3 2(+)	AAGCGCTGCCCCTACACCAA	10-30	146	A9,A10	1
1(-)	CCGGTTCTGGAACCAGATCTT	156-136		A11,C9,D9	2
HOX 4 2(+)	AAGCGCTGCCCCTACACCAA	10-20	140	A9,A13,D9	1
2(-)	CTGAAACCAGATTTTGACCTG	150-130		D10	2
HOX 5 3(+)	GCCCGGACCACCTACACGCG	10-30	146	A4,A6,A7	1
1(-)	CCGGTTCTGGAACCAGATCTT	156-136		B4	2
HOX 6 3(+)	GCCCGGACCACCTACACGCG	10-30	140	A4	1
2(-)	CTGAAACCAGATTTTGACCTG	150-130			2

5

List 5 - a. II. HOX conserved primers

Specific oligonucleotide primers of the HOX genes for RT-PCR reactions

Designation		Sequence 5'-3'	Position [b]	Product size (bp)	Matching the HOX genes:	Ref.
A5	4(+)	CGTTACCTGACCCGCAGAA	70-88	91	A5	2
	3(-)	CATCCTGCCGTTCTGAAACCA	161-141			
A7	5(+)	CACTTCAACCGCTACCTGACC	61-80	95	A7	2
	1(-)	CCGGTTCTGGAACCAGATCTT	156-136			
A10	6(+)	TTCAATATGTACCTTACTCGA	64-84	97	A10	1
	3(-)	CATCCTGCCGTTCTGAAACCA	161-141			
A11	7(+)	GTCTACATTAACAAAGAGAAG	70-90	86	A11	1
	1(-)	CCGGTTCTGGAACCAGATCTT	156-136			
A13	8(+)	CGGGAATACGCCACGAATAAA	51-70	99	A13	1
	2(-)	CTGAAACCAGATTTTGACCTG	150-130			
B1	9(+)	CATTTCAACAAGTACCTGAG	61-80	95	B1	2
	1(-)	CCGGTTCTGGAACCAGATCTT	156-136			
B2	10(+)	AATAAGTACCTGTGCCGGCCA	67-87	83	B2	2
	2(-)	CTGAAACCAGATTTTGACCTG	150-130			
B3	11(+)	AACCGCTACCTGTGCCGGCCT	67-87	89	B3	2
	1(-)	CCGGTTCTGGAACCAGATCTT	156-136			
B5	12(+)	AACCGCTACCTGACCCGGCGA	67-87	89	B5	2
	1(-)	CCGGTTCTGGAACCAGATCTT	156-136			
B6	13(+)	CACTACAATCGCTACCTGACG	61-80	95	B6, B7	2
	1(-)	CCGGTTCTGGAACCAGATCTT	156-136			
B6	13(+)	CACTACAATCGCTACCTGACG	61-80	139		3
	4(-)	AGACGCGCTGAGCAGTTTGC	200-180			
B7	13(+)	CACTACAATCGCTACCTGACG	61-80			2
	5(-)	AGCCGTGTCTTGGCCGGTGGT				
B8	14(+)	TTTAATCCCTATCTGACTCGT	63-83	87	B8	2
	2(-)	CTGAAACCAGATTTTGACCTG	150-130			

B9 15(+)	TACCTCACCAGGGACCGTAGGC	74-94	87	B9	2
3(-)	CATCCTGCCGTTCTGAAACCA	161-141			
C4 16(+)	CGCTACCTGACCCGAAGGAGA	70-90	80	C4	2
2(-)	CTGAAACCAGATTTTGACCTG	150-130			
C5 17(+)	CACTTTAACCGCTACCTCACT	61-81	100	C5	2
2(-)	CATCCTGCCGTTCTGAAACCA	150-130			
C5 18(+)	GGCAAGCGGTCCCGAACCAGT	1-21	161		2
2(-)	CATCCTGCCGTTCTGAAACCA	161-141			
C6(+) 19(+)	TTTCACTTCAATCGCTACCTA	58-78	103	C6	2
3(-)	CATCCTGCCGTTCTGAAACCA	161-141			
C6 19(+)	TTTCACTTCAATCGCTACCTA	58-78			2
6(-)	CCCCGAGAGAGTGGATGTGA				
C8(+) 20(+)	AATCCTTATTTGACACGAAAA	66-86	90	C8	2
1(-)	CCGGTTCTGGAACCAGATCTT	156-136			
C8 21(+)	AGGCGCAGTGGACGGCAA	1-18	156		2
1(-)	CCGGTTCTGGAACCAGATCTT	156-136			
C9(+) 22(+)	GAGTTTCTCTTCAATATGTATTTA	57-77	93	C9	2
2(-)	CTGAAACCAGATTTTGACCTG	150-130			
C12(+) 23(+)	CTGGAGGGCGAGTTTCTGG	46-64	105	C12	1
1(-)	CCGGTTCTGGAACCAGATCTT	156-136			
C12 24(+)	TCTCGGAAGAAGCGCAAG	1-18	183		1
7(-)	CAACAGAAGTCTTTTCTT	183-163			
C13(+) 25(+)	GAGCTAGAGAAGGAATACGCG	43-63	119	C13	1
3(-)	CATCCTGCCGTTCTGAAACCA	161-141			
D3(+) 26(+)	AACCGCTACTTGTGCCGGCCG	67-87	90	D3	2
1(-)	CCGGTTCTGGAACCAGATCTT	156-136			
D4(+) 27(+)	CATTTTAACAGGTATCTGACAA	61-82	96	D4	2
1(-)	CCGGTTCTGGAACCAGATCTT	156-136			
D8(+) 28(+)	AACCCCTATCTGACCAGGAAA	67-87	83	D8	1
2(-)	CTGAAACCAGATTTTGACCTG	150-130			

D10(+) 29(+)	CTCACCCGCGAGCGCCGCCTA	76-96	85	D10	1
3(-)	CATCCTGCCGTTCTGAAACCA	161-141			
D11(+) 30(+)	TTTTCTTTAACGTGTACATA	58-78	103	D11	1
3(-)	CATCCTGCCGTTCTGAAACCA	161-141			

List 6 - b. POU domain containing genes:

I. OCT conserved primers

Designation	Sequence 5'-3'	Position	Product size (bp)	Ref.
OCT 31(+)	CTGAGCTTTAAGAACATGTG	1000-1020	290	4
8(-)	CTGGCGCCGGTTACAGAACCA	1290-1270		

II. OCT specific primers

Designation	Sequence 5'-3'	Position	Product size (bp)	Ref.
OCT-1 32(+)	AGCCCAAGTGCCCTGAATTCT	1077-1097	214	5
8(-)	CTGGCGCCGGTTACAGAACCA	1290-1270		
OCT-2 33(+)	AGCCCCAGCCTGGGTTTC	1021-1038	270	
8(-)	CTGGCGCCGGTTACAGAACCA	1290-1270		

III. LFB specific primers

Designation	Sequence 5'-3'	Position	Product size	Ref.
LFB1(+) 35(+)	TCCGGCGACGAGGGCTCCGAG	208-219	606	6
LFB(-) 11(-)	GCGGTTGGCAAACCAGTTGTA	814-794		
LFB1(+) 36(+)	CTGATTGAAGAGCCCACAGGT	554-574	260	6
LFB(-) 11(-)	GCGGTTGGCAAACCAGTTGTA	814-794		
LFB3(+) 37(+)	TCGGAAGATGACACGGATGAC	420-440	680	7
LFB(-) 11(-)	GCGGTTGGCAAACCAGTTGTA	1100-1079		8
LFB3(+) 38(+)	ATGACAGACAAAAGCAGTCAG	774-794	326	7
LFB(-) 11(-)	GCGGTTGGCAAACCAGTTGTA	1100-1079		8

IV. BRN conserved primers

Designation	Sequence 5'-3'	Position	Product size	Ref.
BRN(+) 34(+)	ACCCTGTATGGCAACGTGTT	1394-1415	212	9
BRN(-) 9(-)	CCCCTTGAGGCTCACCTCGAT	1606-1586		
BRN(+) 34(+)	ACCCTGTATGGCAACGTGTT	1394-1415	288	9
BRN(-) 10(-)	CTGGAGACTGTCTGCCAGCGA	1456-1436		9
		(1682-1662)		9

c. Drosophila like maternally expressed genes

I. CDX conserved primers

Designation	Sequence 5'-3'	Position [a]	Product size	Ref.
CDX 39(+)	GAGCTGGAAAAGGAGTTTCA	43-63	125	10
CDX 12(-)	TCCTTGGCTCTGCGGTTCTG	168-147		11

II. CDX specific primers

Designation	Sequence 5'-3'	Position [a]	Product size	Ref.
CDX 1 40(+)	AAGTCCGAGCTGGCTGCTAAC	91-111	77	10
12(-)	TCCTTGGCTCTGCGGTTCTG	168-147		
CDX 2 41(+)	AAAAGTGAGCTGGCTGCCAC	91-111	77	12
12(-)	TCCTTGGCTCTGCGGTTCTG	168-147		
CDX 3 42(+)	ATCAGGCGGAAAGCTGAATTA	91-111	77	
12(-)	TCCTTGGCTCTGCGGTTCTG	168-147		
CDX 4 43(+)	AGGAAGTCAGAGCTGGCAGTT	91-111	77	11
12(-)	TCCTTGGCTCTGCGGTTCTG	168-147		

III. Bicoid specific primers

Designation	Sequence 5'-3'	Position	Product size	Ref.
Bicoid 44(+)	ATGTCGTCCAGCATGGTGCCC	1110-1130	172	13
13(-)	CGAGTTACACGTGTCCCTATA	1280-1260		
Bicoid 45(+)	AACAGCCTGAATAACTTGAAC	1164-1183	116	13
13(-)	CGAGTTACACGTGTCCCTATA	1280-1260		

d. Drosophila like segmental expressed genes

I. PAX concerved primers

Designation	Sequence 5'-3'	Matces amino acid sequence [b]	Product size	Ref.
PAX 46(+)	GGCTGTGTCAGCAAAATTCT	GCVSKIL	177	14.
14(-)	CAGCCTGTCTCGGATCTCCCA	WEIRDRL		15 16, 17

II. PAX specific primers

Designation		Sequence 5'-3'	Position	Product size	Ref.
PAX 6	46(+)	GGCTGTGTCAGCAAAATTCT	511-530	170	18
	15(-)	CAGCCTGTCTCGGATCTCCCA	681-660		
PAX 7	47(+)	GGCTGTGTCTCCAAGATTCT	341-362	301	19
	16(-)	TTCGCCGTCGTCCTCCTTCTT	642-622		

III. Engrailed specific primers

Designation		Sequence 5'-3'	Position	Product size	Ref.
Engr 1,2	48(+)	AAGATTTGGTTCCAGAACAA	180-200	96	20
	17(-)	GTTGTACAGGCCCTGTGCCAT	275-225		

IV. Goosecoid specific primers

Designation		Sequence 5'-3'	Position	Product size	Ref.
Goosecoid	49(+)	ACCATCTTTACGGAGGAGCA	1330-1350	143	21
	18(-)	GTTCTTGAACCACACCTC	1473-1455		

e. I. LIM domain containing genes conserved primers

Designation		Sequence 5'-3'	Matces amino acid sequence[b]	Product size	Ref.
LIM a	50(+)	GACAAGAAAATTTACTGCAA	DKKLYCK	550-600	22, 23
	19(-)	CGGGCTGCGGTTCTGGAACCA	WFQNQR		24
LIM b	51(+)	GACGGGAAAACCTACTGTAA	DKKLYCK	556-600	
	19(-)	CGGGCTGCGGTTCTGGAACCA	WFQNQR		

II. LIM domain containing genes specific primers

Designation		Sequence 5'-3'	Position	Product size	Ref.
LH 1	52(+)	TTTCACCTCAACTGCTTCAC	682-702	396	25
	20(+)	CTGCGCCAGCTGCTCGCGGAT	1077-1056		
LH 2	52(+)	TTTCACCTCAACTGCTTCAC	630-650	433	26
	21(-)	CATTGTCCGAAGCTGGTGGTG	1062-1042		

f. I. NK genes conserved primers

Designation		Sequence 5'-3'	Position[a]	Product size	Ref.
NK a	53(+)	GTGCTCTTCTCCCAGGCGCAG	16-36	161	27, 28
	22(-)	CTTCATCTTGTAGCGGTG	176-158		29, 30
NK b	53(+)	GTGCTCTTCTCCCAGGCGCAG	16-36	161	
	23(-)	CTTGCACTTGTAGCGACG	176-156		
NK c	54(+)	CTCTTTTCGCAGGGCCCAGGT	18-38	159	
	22(-)	CTTCATCTTGTAGCGGTG	176-156		
NK d	54(+)	CTCTTTTCGCAGGGCCCAGGT	18-38	159	
	22(-)	CTTCATCTTGTAGCGGTG	176-156		
NK e	55(+)	GCACAGGTGTTCGAGCTGGA	33-54	120	
	24(-)	CTGGAACCAGATCTTCACCTG	152-132		

II. NK genes specific primers

Designation		Sequence 5'-3'	Position[a]	Product size	Ref.
NK 2	56(+)	AAACGGAAGCGCCGAGTCCT	3-23	149	27, 30
	24(-)	CTGGAACCAGATCTTCACCTG	152-132		29
NK 3	57(+)	CAGAAGCGCTCCCGAGCTGC	3-23	149	27, 31
	24(-)	CTGGAACCAGATCTTCACCTG	152-132		
NK 6	58(+)	AGAAAACACACGAGACCCACT	3-23	149	29
	24(-)	CTGGAACCAGATCTTCACCTG	152-132		32

g. DLX genes conserved primers

Designation		Sequence 5'-3'	Position[a]	Product size	Ref.
DLX a	59(+)	CGCAAGCCGAGGACCATC	3-23	174	33
	25(-)	CCACATCTTCTTGAACTTGGA	177-159		34
DLX b	59(+)	CGCAAGCCGAGGACCATC	3-23	174	35
	26(-)	CAGGAGCTTCTTATATTTGGA	177-159		36
DLX c	69(+)	CGGAAACCCCGCACCATC	3-23		37
	25(-)	CCACATCTTCTTGAACTTGGA	177-159		38
DLX d	60(+)	CGGAAACCCCGCACCATC	3-23	174	
	26(-)	CAGGAGCTTCTTATATTTGGA	177-159		

f. CNS embryonic genes: EMX, OTX

Designation		Sequence 5'-3'	Position[a]	Product size	Ref.
EMX	61(+)	AAACGCATCCGGACCGCCTT	7-28	150	39
	27(-)	TCGGTTCTGGAACAACACCTT	156-135		
OTX a	62(+)	TATCCGGACATATTCATG	75-94	105	40
	28(-)	CTGTTGGCGGCACTTGGC	162-180		
OTX b	62(+)	TATCCGGACATATTCATG	75-94	105	41
	29(-)	TTTGCGCTTCTTCCATTT	62-180		

h. Miscellaneous HB genes

Designation		Sequence 5'-3'	Position	Product size	Ref.
EVE 1	63(+)	AACCTGCCCGAAACCACCATC	880-901	170	
	30(-)	CAGGGGCAGGTGCGATGG	1641-1621		
EVE 2	63(+)	AACCTGCCCGAAACCACCATC	39-45	147	
	31(-)	GGGGTAGGGCAGGCTTCCGGT	89-83		
GBX 2	64(+)	CGGACTGCCTTCACCAGC	1169-1188	185	42
	32(-)	GGAATTGGCATTGCCTGCCTT	1354-1334		
Meis 1	65(+)	AAAAAGCGTGGCATCTTTCC	1266-1286	192	43
	33(-)	GTCTATCATGGCCTCCACTAT	1457-1436		44
MSX 1	66(+)	CCGCTGGGCCATTTCTCG	760-780	350	45
	21(-)	CGGGCTGCGGTTCTGGAACCA	2190-2170		46
MSX 2	67(+)	AATTCAGAAGATGGAGCGGCG	2190-2170		
	21(-)	CGGGCTGCGGTTCVTGGAACCA			
P HOX	69(+)	GACCAGTTGAACTCTGAGGA	244-264	395	47
	35(-)	AAGCCCCTCGTGTAAACAACA	639-618		48
PROX	70(+)	ATTCAGATGGAGAAGTACGC	2500-2520	263	49
	36(-)	TTCACTATCCAGCTTGCAGAT	2763-2753		

List 7

Designation		Sequence 5'-3'	Position [a]	Matches amino acid motive[b]	Product size (bp)
MAP 1	a(+)	CTGCATAGGGACTTAAAACC	VIb	LHRDLKP	190-210
	a(-)	ACCAACAGACCAAACATC	IX	DVWSVG	
MAP 1	a(+)	CTGCATAGGGACTTAAAACC	VIb	LHRDLKP	190-210
	b(-)	GCCAAGGCTCCAAATGTC	IX	DIWSLG	
MAP 3	a(+)	CTGCATAGGGACTTAAAACC	VIb	LHRDLKP	190-210
	c(-)	GCCCAGGGACCAGATGTC	IX	DIWSLG	
MAP 4	b(+)	ATTCACAGAGACGTCAAGCC	VIb	IHRDVKP	190-210
	a(-)		IX	DVWSVG	
MAP 5	b(+)	ATTCACAGAGACGTCAAGCC	VIb	IHRDVKP	190-210
	b(-)	GCCAAGGCTCCAAATGTC	IX	DIWSLG	
MAP 6	b(+)	ATTCACAGAGACGTCAAGCC	VIb	IHRDVKP	190-210
	c(-)	GCCCAGGGACCAGATGTC	IX	DIWSLG	
MAP 7	c(+)	GTTCACCGGGACATAAAGGG	VIb	VHRDVKG	190-210
	a(-)	ACCAACAGACCAAACATC	IX	DVWSVG	
MAP 8	c(+)	GTTCACCGGGACATAAAGGG	VIb	VHRDVKG	190-210
	b(-)	GCCAAGGCTCCAAATGTC	IX	DIWSLG	
MAP 9	c(+)	GTTCACCGGGACATAAAGGG	VIb	VHRDVKG	190-210
	c(-)	GCCCAGGGACCAGATGTC	IX	DIWSLG	
MAP 10	d(+)	ATTCACAGAGACATCAAGAG	VIb	IHRDIKS	190-210
	a(-)	ACCAACAGACCAAACATC	IX	DVWSVG	
MAP 11	d(+)	ATTCACAGAGACATCAAGAG	VIb	IHRDIKS	190-210
	b(-)	GCCAAGGCTCCAAATGTC	IX	DIWSLG	
MAP 12	d(+)	ATTCACAGAGACATCAAGAG	VIb	IHRDIKS	190-210
	c(-)	GCCCAGGGACCAGATGTC	IX	DIWSLG	

List 8

Designation		Sequence 5'-3'	Position [a]	Matches amino acid motive[b]	Product size (bp)
PTK 1	e(+)	AAAATTGGAGACTTTGGAAT	VII	KIGDFGM	155-165
	e(-)	GCCGAAGCTCCACACGTCGCT	IX	SDVWSFG	
PTK 2	e(+)	AAAATTGGAGACTTTGGAAT	VII	KIGDFGM	155-165
	f(-)	CCCATAGGACCAGACATCACT	IX	SDVWSYG	
PTK 3	e(+)	AAAATTGGAGACTTTGGAAT	VII	KIGDFGM	155-165
	g(-)	TCCAAAGGACCACACATCAGA	IX	SDVWSFG	
PTK 4	e(+)	AAAATTGGAGACTTTGGAAT	VII	KIGDFGM	155-165
	h(-)	CCCGAAGGACCACACGTCCG	IX	SDVWSFG	
PTK 5	f(+)	AAAATTGCAGATTTTGGCCT	VII	KIADFGL	155-165
	e(-)	GCCGAAGCTCCACACGTCGCT	IX	SDVWSFG	
PTK 6	f(+)	AAAATTGCAGATTTTGGCCT	VII	KIADFGL	155-165
	f(-)	CCCATAGGACCAGACATCACT	IX	SDVWSYG	
PTK 7	f(+)	AAAATTGCAGATTTTGGCCT	VII	KIADFGL	155-165
	g(-)	TCCAAAGGACCACACATCAG	IX	SDVWSFG	
PTK 8	f(+)	AAAATTGCAGATTTTGGCCT	VII	KIADFGL	155-165
	h(-)	CCCGAAGGACCACACGTCCGA	IX	SDVWSFG	
PTK 9	g(+)	AAAATAGCAGACTTTGGCTT	VII	KIADFGF	155-165
	e(-)	GCCGAAGCTCCACACGTCGCT	IX	SDVWSFG	
PTK 10	g(+)	AAAATAGCAGACTTTGGCTT	VII	KIADFGF	155-165
	f(-)	CCCATAGGACCAGACATCACT	IX	SDVWSYG	
PTK 11	g(+)	AAAATAGCAGACTTTGGCTT	VII	KIADFGF	155-165
	g(-)	TCCAAAGGACCACACATCAGA	IX	SKVWSFG	
PTK 12	g(+)	AAAATAGCAGACTTTGGCTT	VII	KIADFGF	155-165
	h(-)	CCCGAAGGACCACACGTCCGA	IX	SDVWSFG	

List 8 continued

Designation	Sequence 5'-3'	Position [a]	Matches amino acid motive[b]	Product size (bp)
PTK 13 h(+)	AAGATTTGTGATTTTGGTCT	VII	KICDFGL	155-165
e(-)	GCCGAAGCTCCACACGTCGCT	IX	SDVWSFG	
PTK 14 h(+)	AAGATTTGTGATTTTGGTCT	VII	KICDFGL	155-165
f(-)	CCCATAGGACCAGACATCACT	IX	SDVWSYG	
PTK 15 h(+)	AAGATTTGTGATTTTGGTCT	VII	KICDFGL	155-165
g(-)	TCCAAAGGACCACACATCAG	IX	SDVWSFG	
PTK 16 h(+)	AAGATTTGTGATTTTGGTCT	VII	KICDFGL	155-165
h(-)	CCCGAAGGACCACACGTCCGA	IX	SDVWSFG	

[a] Position is given in protein kinase catalytic unit subdomain nomenclature

[b] Amino acid sequence is given in symbol letters

List 9

Designation		Sequence 5'-3'	Position [a]	Matches amino acid motive[b]	Product size (bp)
STK 1	i(+)	ATCCACCGGGATCTGAACTC	VIb	IHRDLNS	190-210
	i(-)	GCCAAGGGACCACATGTC	IX	DMWSLG	
STK 2	j(+)	GCCCATAGAGACATTAAGCC	VIb	AHRDLKP	290-210
	i(-)	GCCAAGGGACCACATGTC	IX	DMWSLG	
STK 3	k(+)	ACACACCGAGACCTCAAACC	VIb	LHRDLKP	190-210
	i(-)	GCCAAGGGACCACATGTC	IX	DMWSLG	
STK 4	l(+)	GTGCATCGAGATCTAAAGCC	VIb	VHRDLKP	190-210
	i (-)	GCCAAGGGACCACATGTC	IX	DMWSLG	
STK 5	m(+)	CTACACAGGGATATCAAGGC	VIb	LHRDIKP	190-210
	i(-)	GCCAAGGGACCACATGTC	IX	DMWSLG	
STK 6	n(+)	ATACACAGAGATATAAAAGC	VIb	IHRDIKA	190-210
	i(-)	GCCAAGGGACCACATGTC	IX	DMWSLG	
STK 7	o(+)	CTGCACCGTGACCTCAAGAC	VIb	LHRDLKT	290-210
	i(-)	GCCAAGGGACCACATGTC	IX	DMWSLG	
STK 8	i (+)	ATCCACCGGGATCTGAACTC	VIb	IHRDLNS	190-210
	i (-)	GCCCAGCGCCCACATGTC	IX	DMWALG	
STK 9	j(+)	GCCCATAGAGACATTAAGCC	VIb	AHRDLKP	190-210
	j(-)	GCCCAGCGCCCACATGTC	IX	DMWALG	
STK 10	k(+)	ACACACCGAGACCTCAAACC	VIb	LHRDLKP	190-210
	j(-)	GCCCAGCGCCCACATGTC	IX	DMWALG	
STK 11	l(+)	GTGCATCGAGATCTAAAGCC	VIb	VHRDLIP	190-210
	j(-)	GCCCAGCGCCCACATGTC	IX	DMWALG	
STK 12	m(+)	CTACACAGGGATATCAAGGC	VIb	LHRDIKP	190-210
	j(-)	GCCCAGCGCCCACATGTC	IX	DMWALG	
STK 13	n(+)	ATACACAGAGATATAAAAGC	VIb	IHRDIKA	190-210
	j(-)	GCCCAGCGCCCACATGTC	IX	DMWALG	

List 9 (continued)

Designation		Sequence 5'-3'	Position [a]	Matches amino acid motive[b]	Product size (bp)
STK 14	o(+)	CTGCACCGTGACCTCAAGAC	VIb	LHRDLKT	190-210
	j(-)	GCCCAGCGCCCACATGTC	IX	DMWALG	
STK 15	i(+)	ATCCACCGGGATCTGAACTC	VIb	IHRDLNS	190-210
	k(-)	TCCAAGGCTCCAGAGATC	IX	DLWSLG	
STK 16	j(+)	GCCCATAGAGACATTAAGCC	VI b	AHRDLKP	190-210
	k(-)	TCCAAGGCTCCAGAGATC	IX	DLWALG	
STK 17	k(+)	ACACACCGAGACCTCAAACC	VIb	LHRDLKP	190-210
	k(-)	TCCAAGGCTCCAGAGATC	IX	DLWALG	
STK 18	l(+)	GTGCATCGAGATCTAAAGCC	VIb	VHRDLKP	190-210
	k(-)	TCCAAGGCTCCAGAGATC	IX	DLWALG	
STK 19	m(+)	CTACACAGGGATATCAAGGC	VIb	LHRDIKP	190-210
	k(-)	TCCAAGGCTCCAGAGATC	IX	DLWALG	
STK 20	n(+)	ATACACAGAGATATAAAAGC	VIb	IHRDIKA	190-210
	k(-)	TCCAAGGCTCCAGAGATC	IX	DLWALG	
STK 21	o(+)	CTGCACCGTGACCTCAAGAC	VIb	LHRDLKT	190-210
	k(-)	TCCAAGGCTCCAGAGATC	IX	DLWALG	

List 10

Designation		Sequence 5'-3'	Position [a]	Matches amino acid motive[b]	Product size (bp)
PTK 17	p(+)	AAGCTCGGGGACTTTGGGCT	VII	KLGDFGL	155-165
	p(-)	TCCGTAGCTCCAGACATCACT	IX	SDVWSYG	
PTK 18	q(+)	AAGCTCGGGGACTTTGGGCT	VII	KLGDFGL	155-165
	q(-)	AGCAAACATCCAGACGTCACT	IX	SDVWMFA	
PTK 19	p(+)	AAGCTCGGGGACTTTGGGCT	VII	KLGDFGL	155-165
	r(-)	CCCAAAGGACCACACGTCTGA	IX	SDVWSFG	
PTK 20	q(+)	AAGATTTCGGATTTCGGCTT	VII	KISDFGF	155-165
	p(-)	TCCGTAGCTCCAGACATCACT	IX	SDVWSYG	
PTK 21	q(+)	AAGATTTCGGATTTCGGCTT	VII	KISDFGF	155-165
	q(-)	AGCAAACATCCAGACGTCACT	IX	SDVWMFA	
PTK 22	q(+)	AAGATTTCGGATTTCGGCTT	VII	KISDFGF	155-165
	r(-)	CCCAAAGGACCACACGTCTGA	IX	SDVWSFG	
PTK 23	r(+)	AAGATTGGTGATTTTGGCAT	VII	KIGDFGI	155-165
	p(-)	TCCGTAGCTCCAGACATCACT	IX	SDVWSYG	
PTK 25	r(+)	AAGATTGGTGATTTTGGCAT	VII	KIGDFGI	155-165
	r(-)	CCCAAAGGACCACACGTCTGA	IX	SDVWSFG	
PTK 26	e(+)	AAAATTGGAGACTTTGGAAT	VII	KIGDFGI	155-165
	p(-)	TCCGTAGCTCCAGACATCACT	IX	SDVWSYG	
PTK 27	e(+)	AAAATTGGAGACTTTGGAAT	VII	KIGDFGI	155-165
	q(-)	AGCAAACATCCAGACGTCACT	IX	SDVWMFA	
PTK 28	e(+)	AAAATTGGAGACTTTGGAAT	VII	KIGDFGI	155-165
	r(-)	CCCAAAGGACCACACGTCTGA	IX	SDVWSFG	
PTK 30	g(+)	AAAATAGCAGACTTTGGCTT	VII	KIADFGF	155-165
	p(-)	TCCGTAGCTCCAGACATCACT	IX	SDVWSYG	

List 10 (continued)

Designation	Sequence 5'-3'	Position [a]	Matches amino acid motive[b]	Product size (bp)
PTK 31 g(+)	AAAATAGCAGACTTTGGCTT	VII	KIADFGF	155-165
q(-)	AGCAAACATCCAGACGTCACT	IX	SDVWMFA	
PTK 32 g(+)	AAAATAGCAGACTTTGGCTT	VII	KIADFGF	155-165
r(-)	CCCAAAGGACCACACGTCTGA	IX	SDVWSFG	

[a] Position is giving in protein kinase catalytic unit subdomain nomenclature

[b] Amino acids are given by single letter abbreviations.

List 11 - Serine Threonine kinase

Primer	Sequence 5'-3'	Conserved Domain	Best matches the gene
240 (+)	ATC CAC CGC GAC CTG AAG CC	IHRDLKP	PKC γ
241 (+)	ATT TAC AGG GAC CTC AAG CT		PKC
242 (+)	ATC TAC AGA GAT CTA AAA CC	IYRDLKP	PKC β
243 (+)	ATT TAC CGT GAC CTG AAA CT	IYRDLKC	PKC
244 (+)	ATC TAC AGG GAT TTG AAA CT	IYRDLKC	PKC
245 (+)	ATC CAC CGA GAT CTC AAG TC	IHRDLKS	A RAF
246 (+)	GTG CAC AGG GAC CTC AAG AG	VHRDLKS	TGFβ 2 recep
247 (+)	GCC CAC AGG GAC TTC AAA AG	AHRDFKS	Activin recep
248 (-)	GCC CAT GGC GTA CAT GTC	DMYAMG	Activin receptor
249 (-)	AGC CAT GGA GTA GAC ATC	DVYSMA	TGFβ2 recept
250 (-)	CAG CGA AAA GCT GTA CAC ATC	DVYSFS	daf 1 receptor
251 (-)	GCC CAG GGC CCA CCA GTC	DWWALG	PKC, cAMP α
252 (-)	CCC AAA AGA CCA CCA GTC	DWWSFG	PKC δ Human
253 (-)	GCC AAA CGC CCA CCA GTC	DWWAFG	PKC
254 (-)	CCC GTA GGC ATA GAC GTC	DVYAYG	A RAF
255 (-)	AAT TCC AAA TGC ATA TAC ATC	DVYAFG	B RAF

List 12 - Tyrosine kinase

Primer	Sequence 5'-3'	ConservedDomain	Best matches the gene
400 (+)	AAA GTC TCA GAC TTT GGC CT	KVSDFGL	HEK2
401 (+)	AAA GTC TCT GAT TTC GGC CT	KVSDFGL	Eph
402 (+)	AAA GTG TCA GAT TTT GGC AT	KVSDFGI	Eph related
403 (+)	AAG ATC ACA GAC AAT GCC CT	KITDNAL	RYK
404 (+)	CAG GTG GCA GAT TTT GGT GT	QVADFGV	HER 3
405 (+)	AAG GTG GCT GAC TTT GGT TT	KVADFGL	MET receptor
408 (-)	TCC AAA GGC CCA CAC GTC ACT	SDVWAFG	RYK
409 (-)	CCC GAA GGA CCA CAC ATC GCT	SDVWSFG	AXL

List 13 - Tyrosine phosphatase

Primer	Sequence 5'-3'	Conserved Domain	matches amino acid motive
705 (+)	GAC TTC TGG GAG ATG GTG TGG	PTPase	DFWEMVW
706 (+)	GAC TTC TGG AGG ATG ATC TGG	PTPase	DFWRMIW
708 (+)	AAG TGT GAT CAG TAC TGG CC	PTPase	KCDQYWP
709 (+)	AAA TGT GAC CAT TAC TGG CC	PTPase	KCDHYWP
710 (+)	AAG TGT CAC CAA TAT TGG CC	PTPase	KCHQYWP
711 (+)	AAA TGC TGC AAA TAC TGG CC	PTPase	KCCKYWP
721 (-)	AGC ACC AGC ACT GCA ATG TAC	PTPase	VHCSAGA
722 (-)	CAC GCC TGC GCT ACA GTG GAC	PTPase	VHCSAGI
723 (-)	GCC CAC GCC CGC GCT ACA GTG	PTPase	HCSAGVG
724 (-)	GCC AAT CCC TGC ACT GCA GTG	PTPase	HCSAGIG
725 (-)	TTG CAC CAA GTA GTT CCT CTG	PTPase	QRNYLVQ
726 (-)	CTG TTC CTC TGT CTG CAC CAT	PTPase	MVQTEEQ
727 (-)	GTA CTG TTC CTC TGT CTG GAC	PTPase	VQTEEQY
728 (-)	CTT GTA GCA GAA TTC ATA CTG CTC	PTPase	QYEFCYK

Average distance between sense primers 705-706 and antisens primers 725-729 of Tyrosine phosphatase is 450-480 bp.
Average distance between sense primers 708-711 and antisense primers 721-724 of Tyrosine phosphatase enzymes is 270 bp.

List 14 - Serine Threonine Phosphatase:

Primer	Sequence 5'-3'	Conserved Domain
740 (+)	CTC TTT CTG GGG GAC TAT GTG	PSTPase
741 (+)	TTT AAT GGT GAC TTT GTG GAC	PSTPase
742 (+)	ATG GGT GAT TTT GTA GAC AG	PSTPase
743 (+)	TAC CTC TTC TTA GGG GAC TAT GT	PSTPase
744 (+)	CTC TTC ATG GGG GAC TTT GTG	PSTPase
745 (+)	TTT ATG GGA GAT TAT GTT GAC	PSTPase
746 (-)	GTC AGA CCA CAG CAG GTC ACA CA	PSTPase
747 (-)	ATC TGA CCA TAA CAA ATC ACA CA	PSTPase
748 (-)	ATC TGA CCA AAC CAG ATC ACA AA	PSTPase
749 (-)	ATC TGA CCA GAG CAG GTC ACA CA	PSTPase
750 (-)	ATC CGA CCA GAG GAG GTC GCA CA	PSTPase
751 (-)	GTC AGA CCA CAA AAG ATC ACA AA	PSTPase

Average distance between sense and antisense primer of PCR products of the Serine Threonine Phosphatase is 350-370 bp.

List 15 - P450 enzymes

Primer	Sequence 5'-3'	Conserved Domain	amino acid motiv	matches the gene
800 (+)	CTG ACC ATG TGC ATC AAG GAG	Steroid bind	LTMCIKE	4F3 4B1
801 (+)	GAG ACC TTG CGG CTC TAC CCT	Steroid bind	ETLRLYP	27, 11B2
802 (+)	ATC AAG GAG ACA CTA AGA CT	Steroid bind	IKETLRL	11A1,51
803 (+)	GTG GTC AAG GAA GTG CTA AGA	Steroid bind	VVKEVLR	25 1 α
804 (+)	ATT AAG GAG GCA CTG AGG CT	Steroid bind	IKEALRL	4A, 11
806 (+)	ATC AAG GAG ACC CTT CGA CT	Steroid bind	IKETLRL	26
809 (+)	ATG CCC TAC ATG GAT GCT GT		MPYMDAV	2
810 (+)	ATG CCC TAC ACC AAT GCT GT		MPYTNAV	2
811 (+)	ATG CCC TAC ATG GAG GCA GT		MPYMEAV	2
812 (+)	ATG CCT TAC ACA GAC GCG GT		MPYTDAV	2
820 (-)	TCG CTT TCC GGT GGA GAA AGG	Iron binding	PFSTGKR	2C10, 2F1, 2A6
821 (-)	ATG ACG CCC AGC TCC AAA TGG	Iron binding	PFGAGRH	51
822 (-)	CCT AAG GCC TCC TCC AAA TGG	Iron binding	PFGGGLR	26
823 (-)	GCG CAT GCC AAA GCC AAA	Iron binding	FGFGMR	11B2,25D3 1α
824 (-)	CTG CCG CAC ACC CCA GCC AAA	Iron binding	FGWGVRQ	11B1, 11A1
825 (-)	CTG GCG CAC GCC ATA GCC GAA	Iron binding	FGYGVRA	27, 11B2
826 (-)	GTT CCT GGG CCC TGC TGA GAA	Iron binding	FSAGPRN	4F3, 4B1
827 (-)	CAG GAC TTC CCC GAT ACA CCG	Iron binding	RCIGEVL	1A1, 1B1
828 (-)	CAT TCG AGC CAG GCC TTC TCC	Iron binding	GEGLARM	2C10,2F1,2A6
829 (-)	AGC CAG GCG TCT GCC CAT GCA	Iron binding	CLGRRLA	11B2,25D3 1α
830 (-)	AGC GAT GCG CCG TCC CAG ACA	Iron binding	CLGRRIA	11B1, 11A1
831 (-)	GGC AAA GTT CTC CCC AAT ACA	Iron binding	CIGENFA	51
832 (-)	TGC AAA TTC TTT GCC TAC ACA	Iron binding	CVGKEFA	26
835 (+)	CGG TGA AGA AAC CGT ATC CAT		(1501-1521)	19
836 (-)	GGG CTT AAT TCA CAG CAA GGG		(1801-1821)	19

Average distance between sense primers 800-806 and antisense primers of P450 enzymes is 250-300 bp.
Average distance between sense primers 809-812 and antisense primers of P450 enzymes is 350 bp.

List 16 - Steroid receptor super family

Primer	Sequence 5'-3'	Conserved Domain	matches the genes
500 (+)	AGT TGC GAA GGC TGC AAA GGG TT	DNA binding	RXRγ
501 (+)	AGC TGC GAG GGG TGC AAG GGC TT	DNA binding	RXRα
502 (+)	ACA TGT GAA GGC TGC AAG GGA TT	DNA binding	RORβ
503 (+)	AGT TGT GAA GGT TGC AAA GGT TT	DNA binding	TR4
504 (+)	ACC TGT GAA GGC TGC AAA GGC TT	DNA binding	Vitamin
505 (+)	TCG TGT GAG GGA TGT AAG GCC TT	DNA binding	Esrogen
506 (+)	ACC TGT GGG AGC TGT AAG GTC TT	DNA binding	Progesterone
507 (+)	ACT TGT GGC AGC TGT AAA GTT TT	DNA binding	GCR α
511 (-)	AGC TTC CTT GGA CAT GCC CAC	DNA binding	RARγ
512 (-)	GCC ACA CTT CAC CAT TCC CAC	DNA binding	Estrogen recep
513 (-)	GTC ATC TAG AAC CAA GTC CAT	DNA binding	Thyroid recept
514 (-)	AGC TTC CAT CTT CAT GCC CAT	DNA binding	TR4/RXRγ
515 (-)	CAC TTC TTT GAC CAT CCC AAC	DNA binding	TINUR=NGFI
516 (-)	GAA TTC CTT CAT CAT GCC GAT	DNA binding	Vitamin D rece
517 (-)	AGC CTC CAG GTT CAT TCC AGC	DNA binding	GCRα2
518 (-)	ACC TCC AAG GAC CAT GCC AGC	DNA binding	Progesterone
519 (-)	CGC CTC CCT CTT CAT GCC AGT	DNA binding	RXRβ (-)
521 (-)	GGC TCC CAG AGT CAT CCC TGC	DNA binding	Androgen rece

Average distance between sense and antisense primer of PCR products of the Steroid receptor super family is 150 bp.

List 17 - Cadherin super family

Primer	Sequence 5'-3'	Conserved Domain	matches amino acid motiv	Best matches the gene
550 (+)	CCT GCC ATT CTG GGG ATT CTT	trans membr	PAILGIL	E-Cad
551 (-)	GTT CAA GTA GTC ATA GTC CTG	intra cytopla		E-Cad
553 (+)	GGT GCC ATC ATT GCC ATC CTG	trans membr	GAIIAIL	N-Cad
554 (-)	GTT CAG GTA ATC ATA GTC CTG	intra cytopla	QDYDYLN	N-Cad
555 (+)	GGT TTC ATC CTC CCT GTG CTG	trans membr	GFILPVL	P-Cad
556 (-)	GTT CAG ATA ATC GTA ATC TTG	intra cytopla	QDYDYLN	P-Cad
557 (+)	GGG GCA CTG ATC GCC ATC CTT	trans membr	GALIAIL	OB-Cad
558 (-)	CTG TAG ATA GTC GTA GTC CAG	intra cytopla	LDYDYLQ	OB-Cad
559 (+)	GTC TGT CCC AGC GGC AGG TTT	trans membr	VCPSGRF	Fat Cad
560 (-)	GTT CAA ATA GAT GGA GTC CAT	intra cytopla	MDSIYLN	Fat Cad
561 (+)	GGA GCC TTA ATC GCT ATT CTT	trans membr	GALIAIL	Cad-14
562 (-)	TCC AAG GTA GTG ATA ATC CTG	intra cytopla	QDYDYLG	Cad-14
563 (+)	GGC GCA CTG GTC ATC GTG CTG	trans membr	GALVIVL	Cad-15
564 (-)	TCT GAG GTA GTC GTA GTC CTG	intra cytopla	QDYDYLL	Cad-15
570 (+)	ACT GGG GCC TTG ATT GCC ATC CT	trans membr		Cad-6,11, 12
571 (+)	ATC CAG GCA GTG GTA GCC ATC TT	trans membr		Cad 5
572 (+)	ATG GGC GCC TTA ATT GCC ATA TT	trans membr		Cad 8
573 (-)	AGG TCC CCA GTC ACT AAG GTA	intra cytopla		Cad 6,8
574 (-)	TCC CCA GTC GTT AAG GAA GTC	intra cytopla		Cad 5
575 (-)	ATC GTA GTC CAA GAC TGA ATC	intra cytopla		Cad 11
576 (-)	TCC CCA GTC TGT CAG ATA GTC	intra cytopla		Cad 12
580 (+)	CTT GGA AAG TGG GCC ATC CTT	trans membr	LKGWAIL	DGII
581 (-)	TTG TCG TTC ACT GCA ACA ACC	intra cytopla	GCCCSER	DGII
582 (+)	CTT GGA AGA TGG GCT ATT CTT	trans membr	LRGWAIL	Desmocol 1
583 (-)	CTG CCG ATC ACT GCA GCA ACC	intra cytopla	GCCSDRQ	Desmocol 1

List 17 (continued)

Primer	Sequence 5`-3`	Conserved Domain	matches amino acid motive	Best matches the gene
584 (+)	CTT GGA AAA TGG GCA ATC CTT	trans membr	LKGWAIL	Desm novel
585 (-)	CTG CTT TTC ACT GCA GCA GCC	intra cytopla	GCCSERQ	Desm novel
586 (+)	GGT CCT GCT GGC ATT GGA CTC	trans membr	GPAGIGL	Dsg1/ DGI
587 (-)	TCC AAT GAA GCT ACA ACA ACC	intra cytopla	GCCSFIG	Dsg1/ DGI
588 (+)	GGA CCC GCA GCA ATT GCG CTC	trans membr	GPAAIAL	Dsg 2
589 (-)	TTC AAT AAA ACT GCA ACA ACC	intra cytopla	GCCSFIE	Dsg 2
590 (+)	AAG CAG CGT GGC AAC ATT CTC	trans membr	KQRGNIL	ProtoCad42
591 (-)	GTA GCT GTA GTC TGA GTA CTG	intra cytopla	QYSDYSY	ProtoCad42
592 (+)	AAA CAG AGA CTC AGT ATT GTC	trans membr		ProtoCaH7
593 (-)	ACA GCT GTA CTC CGA GCA GTG	intra cytopla	H SEYS	ProtoCaH7

Average distance between sense and antisense primer of PCR products of the Cadherin super family is 450-480 bp.

List 18 - Homeobox transcription factors

Primer	Sequence 5'-3'	Conserved Domain	matches the genes
320 (+)	CAG TCG GCC ATC TGC CGG TT	POU	mPOU, Retinal
321 (+)	CAA ACA ACC ATC TGC CGA TTT GAA	POU	Pit-1, OCT-1
322 (+)	CAA ACG ACC ATC TGC CGC TTT GAG	POU	OTF-3 (+)
323 (+)	CAG ACC ACC ATC TGC AGG TTCGAG	POU	BRN-3,BRN-4, OTF-1
324 (+)	CAG ACC ACC ATC TGC CGC TTC GAG	POU	BRN-1, OTF-2, OCT6
326 (-)	CTG CCT CTG GTT GCA GAA CCA	HD	Retina, BRN-3, OTF1
327 (-)	CTG GCG CCG GTT ACA GAA CCA	HD	OCT-6,OTF-2,Pit1, BRN-1, mPOU
328 (-)	TTG GCG CCG ATT GCA GAA CCA	HD	OTF3A+B,BRN4,OCT1
330 (+)	AAA AAG AAG ACG CGC ACC GTC	KKKTRTV	Nkx 5.1, 5.2
331 (+)	AAG AGA AAA CAC ACG AGA CCC	KRKHTRP	Nkx 6.1
332 (+)	AAG CGC TCC CGA GCT GCC TTC	KRSRAAF	Nkx 3.1
333 (+)	CGG AGG AAG CCG CGC GTG CTC	RRKPRVL	HCSX
334 (+)	AAG CGG AAG CGC CGA GTG CTC	KRKRRVL	Nkx 2.2
336 (-)	CTG GAA CCA GAT CTT GAC CTG	QVKIWFQ	Nkx 5.1
337 (-)	CTG GAA CCA AGT CTT AAC CTG	QVHTWFQ	Nkx 5.2
338 (-)	CTG GAA CCA GAT TTT CAC CTG		HCSX,Nkx 3.1, 6.1

Average distance between sense and antisense primers of the POU domain transcription factor is 300 bp.

Average distance between sense and antisense primers of the Nkx transcription factor is 150 bp.

List 19 - House keeping genes

Primer designation	Sequence 5'-3`	Position in the gene	matches the gene
303 (+)	TAC AAT GAG CTG CGT GTG GCT	aa (91-97)	β- Actin
301 (+)	ATG GAG TCC TGT GGC ATC CAC	aa (268-274)	β- Actin
300 (-)	GCC GAT CCA CAC GGA GTA CTT	aa (341-335)	β- Actin
307 (+)	TTC ATT GAC CTC AAC TAC ATG	(105-125)	GAPDH
305 (+)	GCT GCC AAG GCT GTG GGC AAG	(630-650)	GAPDH
304 (-)	GTG CTC AGT GTA GCC CAG GAT	(833-813)	GAPDH

Levels of expression lower than any of the 4 house keeping genes' PCR products are scored as 1, Levels of expression equal or higher than the highest of the 4 products are scored as 3. Any levels in between are scored as 2.

PCR reactions were produced with the above listed primers.
The following tables summerize the results of the PCR expression levels in several tissues. They include primers designation and their composition in the PCR reactions.

REFERENCES

1. Acampora, D., *et al., Nucl. Acids. Res.*, 17:10385-10402, 1989.

2. Boncinelli, E., *et al., Genome,* 31:745-756, 1989.

3. Giampaolo, A. *et al,* Differentiation, 40: 191-197, 1989. .

4. Petryniak, *et al., Proc Natl. Acad. Sci.*, USA, 87:1100-1104, 1990.

5. Sturm, R.A., *et al., Genes Dev.,* 2:1582-1599, 1988.

6. Chourd. T. *et al., Genes Dev.,* 6:715-729, 1992.

7. Bach. I., *et al., Nucl. Acids. Res.,* 19:3553-3559, 1992.

8. Abbott, C., *et al., Genomics,* 8:165-167, 1990.

9. Hara, Y., *et al., Proc. Natl. Acad. Sci.*, USA. **89**:3280-3284, 1992.

10. Hu, Y., *et al., J. Biol. Chem.*, **268**:27214-27225, 1993.

11. Gamer, L.W., and Wright, C.V.E., *Mech. Dev.*, **43**:71-81, 1993.

12. James, R., *et al., J. Biol. Chem.*, **269**:15229-15237, 1994.

13. Semina, E.V., *et al., Genet.*, **14**:392- , 1996.

14. Macina, R.A., *et al., Development*, **119**:419-431, 1993.

15. Plachov, D., *et la., Development*, **110**:643-651, 1990.

16. Adams, B., *et al., Genes Dev.*, **6**:1589-1607, 1992.

17. Rothenpieler, U.W., and Dressler, G.R., *Development*, **119**:711-720, 1993.

18. Ton, C.C.T., *et al., Cell*, **67**:1059-1074, 1991.

19. Schafer, B.W., *et al., Nucl. Acids Res.*, **22**:5474-5482, 1994.

20. Poole, S.J., *et al., Genomics*, **4**:225-231, 1989.

21. Bloom, M., *et al., Genomics*, **21**:388-393, 1994.

22. German, M.S., *et al., Genes Dev.*, **6**:2165-2176, 1992.

23. Freyd, G., *et al., Nature*, **344**:876-881, 1990.

24. Cohen, B., *et al., Genes Dev.*, **6**:715-729, 1992.

25. Dong, W.F., *et al., DNA Cell Biol.*, **6**:671-678, 1997.

26. Xu, Y., *et al., Proc. Natl. Acad. Sci.*, USA, **90**:227-231, 1993.

27. Kim, Y., and Nirenberg, M., *Proc. Natl. Acad. Sci.* USA, **86**:7716-7720, 1989.

28. Guazzi, S., *et al., EMBO J.*, **9**:3631-3639, 1990.

29. Lints, T.J., *et al., Development*, **119**:419-431, 1993.

30. Price, M., *et al., Neuron*, **8**:241-255, 1992.

31. He, W.W., *et al., Genomics*, **43**:69-77, 1997.
32. Inoue, H., *et al., Genomics*, **40**:367-370, 1997.

33. Simeone, A., *et al., Proc. Natl. Acad. Sci.* USA, **91**:2250-2254, 1994.

34. Quinn, L.M., *et al., Gene*, **187**:55-61, 1997.

35. Cohen, S.M., *et al., Nature*, **338**:432-434, 1989.

36. Selski, D.J., *et al., Gene*, **132**:301-303, 1993.

37. Scherer, S.W., *et al., Mamm. Genome*, **6**:310-311, 1995.

38. Porteus, M.H., *et al., Neuron*, **7**:221-229, 1992.

39. Dalton, D., *et al., Genes, Dev.*, **3**:1940-1956, 1989.

40. Finkelstein, R., *et al., Genes Dev.*, **4**:1516-1527, 1990.

41. Simeone, A., *et al., Neuron*, **13**:83-101, 1994.

42. Lin, X., *et al., Genomics*, **31**:335-342, 1996.

43. Moskow, J.J., *et al., Moll. Cell, Biol.*, **15**:5434-5443, 1995.

44. Steelman, S., *et al., Genome Res.*, **7**:142-156, 1997.

45. Monaghan, A.P., *et al., Development*, **112**:1053-1062, 1991.

46. Hewitt, J.E., *et al., Genomics*, **11**:670-678, 1991.

47. Gruenberg, D.A., *et al., Science*, **257**:1089-1091, 1992.

48. Cserjesi, P., *et al., Development*, **115**:1087-1101, 1992.

49. Zinovieva, R.D., *et al., Genomics*, **35**:517-522, 1996.

Table 1

Expression pattern of conserved HOX genes

Hoxgene	Placenta	*S. nigra*	Colon normal	Colon tumor
A5	0	0	0	0.5
A7	1	1	1	1
A11	1	1	0.5	1
A13	1	0	1	1
B1	0	0	1	1
B2	0	0	0	0
B3	0	0	0	0
B6.7	1	1	1	1
C5	0	0	0	0
C8	0	0	0	0
C12	0	0	0.5	1
C13	0	0	0	0
D3	0	0	1	1
D10	0	0	0	0
D11	0	0	0	0

Table 2

	Placenta	*S. nigra*	Colon normal	Colon tumor
MSX2	1	0	0	0
OCT	1	0	0	0
LFB1	1	1	0	0
EN	0	1	0	0
BICOID	1	0	0	0
CDX	0.5	1	1	1
PAX	0	0	0	0

Table 3a

Expression pattern of specific HOX genes

Hoxgene	Placenta	Adrenal	Fetal kidney
A5	0	1	1
A7	1	1	1
A11	1	1	1
A13	1	1	0
B1	0	0	1
B2	0	0	0
B3	0	0	0
B6,7	1	1	1
C5	0	0	
C8	0	0	
C12	0	0	0
C13	0	0	0
D3		1	1
D10	0	0	0
D11	0	0	0

Table 3b

Expression pattern of dispersed homeobox genes

	Placenta	Adrenal	Fetal kidney
MSX2	1	0	0
OCT	1	0	1
LFB1	1	1	1
EN	0	1	0.5
BICOID	0.5	0	0
CDX	0.5	1	1
PAX	0	0	1

Table 4: Expression pattern of HOX genes

HOX gene	Placenta	*S. nigra*	Colon normal	Colon tumor	F. kidney	A. kidney	Adrenal	F. brain	A. brain
HOX 1	2	2	3	3	N.C	N.C	3	2	2
HOX 2	0	0	0	0	N.C	N.C	1	0	0
HOX 3	2	0	2	2	N.C	N.C	2	0	0
HOX 4	3	0	2	2	N.C	N.C	2	1	0
HOX 5	2	2	3	3	N.C	N.C	3	2	2
HOX 6	2	0	2	2	N.C	N.C	2	0	0
A5	0	0	0	1	2	2	2	N.C	N.C
A7	2	2	2	2	3	2	3	N.C	N.C
A10	0	0	0	1	2	1	1	N.C	N.C
A11	2	2	1	2	2	0	2	N.C	N.C
A13	2	0	2	2	0	2	2	N.C	N.C
B1	0	0	2	2	2	2	0	N.C	N.C
B2	0	0	0	0	2	0	0	N.C	N.C
B3	0	0	0	0	0	2	0	N.C	N.C
B5	2	2	3	3	2	2	2	N.C	N.C
B6	3	2	3	3	3	3	2	N.C	N.C
B7	3	2	3	3	3	3	2	N.C	N.C
B8	1	0	1	2	2	2	2	N.C	N.C
B9	2	0	2	2	2	2	1	N.C	N.C
C4	0	0	1	2	0	0	0	N.C	N.C
C5	0	0	0	0	0	1	2	N.C	N.C
C6	0	0	0	0	2	2	2	N.C	N.C
C8	0	0	0	0	2	2	2	N.C	N.C
C9	0	0	0	1	2	2	2	N.C	N.C
C12	0	0	1	2	0	0	0	N.C	N.C
C13	0	0	0	0	0	0	0	N.C	N.C
D3	0	0	2	2	2	2	2	N.C	N.C
D4	N.C	0	N.C	N.C	2	2	2	N.C	N.C
D8	2	2	2	2	2	3	2	N.C	N.C
D10	0	0	0	0	0	2	0	N.C	N.C
D11	0	0	0	0	0	0	0	N.C	N.C

Table 5: Expression pattern of dispersed homeobox genes

Dispersed homeobox	Placenta	*S. nigra*	Colon normal	Colon tumor	F. kidney	A. kidney	Adrenal	F. brain	A. brain
OCT	2	0	0	0	2	1	1	2	1
OCT 1*	2	2	3	1	2	2	2	2	3
OCT 2	2	1	0	0	0	1	1	0	0
BRN a	0	1	0	0	0	0	0	N.C	2
BRN b	0	1	0	0	0	0	0	N.C	1
LFB 1	0	0	0	0	0	0	0	0	0
LFB 3	3	2	0	0	2	2	2	2	2
CDX	2	3	3	3	3	3	3	0	0
CDX 1	0	0	0	0	0	0	0	0	0
CDX 2	0	0	2	2	0	0	0	0	0
CDX 3	0	0	0	0	0	0	0	0	0
CDX 4	2	2	0	0	2	2	2	0	0
Bicoid	2	0	0	0	0	1	0	0	0
PAX	0	0	0	0	2	1	0	N.C	0
PAX 6	0	0	0	0	0	0	0	N.C	0
PAX 7	0	0	0	0	0	0	0	N.C	0
Engrailed	1	3	0	0	1	2	3	2	3
Goosecoid	0	0	0	0	0	0	0	0	0
LIM a	3	2	2	2	N.C	N.C	0	2	2
LIM b	3	2	2	2	N.C	N.C	0	2	1
LH 1	0	0	0	0	0	0	0	0	0
LH 2	0	0	0	0	0	0	0	0	0
NK a	0	0	0	0	0	0	0	0	0
NK b	0	0	0	0	0	0	0	0	0
NK c	1	1	2	2	2	2	2	0	0
NK d	1	1	2	2	2	2	2	0	0
NK e	0	0	0	0	0	N.C	0	0	0
NK 2	0	1	0	0	0	N.C	0	0	0
NK 3	0	1	0	0	2	N.C	0	0	0
NK 6	0	0	0	0	0	N.C	0	0	0
DLX a	0	0	0	0	0	N.C	0	0	0
DLX b	2	1	0	0	0	N.C	0	0	0
DLX c	0	0	0	0	0	N.C	0	2	2
DLX d	0	0	0	0	0	N.C	0	0	0

Table 5 continued

Dispersed homeobox	Placenta	*S. nigra*	Colon normal	Colon tumor	F. kidney	A. kidney	Adrenal	F. brain	A. brain
EMX	0	0	0	0	0	0	0	0	0
OTX a	0	0	0	0	0	0	0	0	0
OTX b	0	0	0	0	0	0	0	0	0
EVE 1	0	0	0	0	0	0	0	0	0
EVE 2	0	0	0	0	0	0	0	0	0
GBX	1	1	0	0	0	N.C	0	2	2
MEI	2	2	0	0	2	N.C	2	2	2
MSX 1	0	0	0	0	0	0	0	0	0
MSX 2	3	0	0	0	0	0	0	0	0
PBX*	3	3	2	3	3	3	3	3	2
P HOX	1	1	0	0	0	N.C	0	0	0
PROX	0	2	2	2	2	N.C	2	2	3

* Genes that serves as reference to the PCR HB expression.

N.C = not checked.

The homeobox proximity indices for the pairs of tissues tested are summarized in Table 6.

Table 6

	Placenta	*S. nigra*	Colon N	Colon T	F. kidney	Adrenal	F. brain	A. brain
Placenta	1	0.51	0.54	0.46	0.47	0.49	0.41	0.3
S. nigra	0.51	1	0.39	0.35	0.42	0.44	0.56	0.58
Colon N	0.54	0.39	1	0.83	0.46	0.51	0.33	0.29
Colon T	0.46	0.35	0.83	1	0.56	0.57	0.36	0.24
F. kidney	0.47	0.42	0.46	0.56	1	0.68	N.C	0.36
Adrenal l	0.49	0.44	0.51	0.57	0.68	1	0.40	0.36
F. Brain	0.41	0.56	0.33	0.36	N.C	0.40	1	0.71
A. Brain	0.3	0.58	0.29	0.24	0.36	0.36	0.71	1

Table 7

Scoring of the expression pattern of MAP conserved region in various cells and organs

MAP designation	Placenta	Amniotic culture 1	Colon normal	Colon tumor	F. kidney	Adrenal	F. brain	A. brain
MAP 1	-	-	1	-	2	2	2	2
MAP 2	2	2	3	2	3	2	3	2
MAP 3	2	1	-	-	-	1	-	2
MAP 4	-	-	1	-	1	2	2	2
MAP 5	2	-	3	1	2	2	3	2
MAP 6	2	2	2	-	2	2	2	2
MAP 7	-	-	-	-	-	-	-	-
MAP 8	-	-	-	-	-	-	-	-
MAP 9	3	2	2	2	2	3	2	3
MAP 10	-	-	-	-	-	-	-	-
MAP 11	-	-	-	-	1	-	-	1
MAP 12	3	3	3	2	3	3	3	3

Table 8

Scoring of the expression pattern of PTK conserved region in various cells and organs

PTK designation	Placenta	Amniotic culture 1	Colon normal	Colon tumor	F. kidney	Adrenal	F. brain	A. brain
PTK 1	-	-	1	1	1	1	-	-
PTK 2	-	-	-	1	-	-	-	-
PTK 3	3	-	2	2	1	2	1	-
PTK 4	-	1	1	2	-	-	-	-
PTK 5	3	2	-	2	2	1	2	-
PTK 6	3	2	2	2	3	3	2	3
PTK 7	3	2	3	3	3	3	2	3
PTK 8	2	1	2	2	2	2	1	2
PTK 9	-	3	2	3	-	-	-	-
PTK 10	2	1	1	-	-	1	1	-
PTK 11	2	1	2	2	2	1	1	2
PTK 12	-	-	1	-	-	-	-	-
PTK 13	1	-	-	-	1	-	-	-
PTK 14	-	-	-	-	-	-	2	2
PTK 15	-	-	-	-	-	-	-	-
PTK 16	-	-	-	-	-	-	-	-

Table 9

	Placenta	Amniotic culture 1	Colon normal	Colon tumor	F. kidney	Adrenal	F. brain	A. brain
Placenta	1	0.50	0.56	0.50	0.65	0.68	0.48	0.57
Amniotic cells	0.50	1	0.57	0.56	0.51	0.50	0.58	0.42
Colon N	0.56	0.57	1	0.60	0.65	0.67	0.61	0.56
Colon T	0.50	0.56	0.60	1	0.51	0.49	0.40	0.38
F. kidney	0.65	0.51	0.65	0.51	1	0.73	0.68	0.65
Adrenal	0.68	0.50	0.67	0.49	0.73	1	0.67	0.68
F. Brain	0.48	0.58	0.61	0.40	0.68	0.67	1	0.62
A Brain	0.57	0.42	0.56	0.38	0.65	0.68	0.62	1

Table 10

Culture Number	Karotype	Sex chromosomes	Gestational age (weeks)	Maternal age (years)
1	Normal	XY	18-19 weeks	38 years
2	Normal	XX	20 weeks	34 years
3	Normal	XX	20 weeks	26 years
4	Trisomy 21	XY	17 weeks	37 years
5	Normal	XX	20 weeks	39 years
6	Normal	XY	19 weeks	26 years
7	T 1:13	XY	17 weeks	36 years
8	Trisomy 21	XY	23 weeks	34 years

Table 11

The number of genes of each type expressed in placenta and amnion

HOX gene	Placenta	Amnion	Dispersed homeobox	Placenta	Amnion
HOX 1	2	3	OCT	2	0
HOX 2	0	0	OCT 1*	2	2
HOX 3	2	2	OCT 2	2	N.C
HOX 4	3	2	BRN a	0	1
HOX 5	2	3	BRN b	0	2
HOX 6	2	2	LFB 1	0	0
			LFB 2	3	0
A5	0	0			
A7	2	2	CDX	2	2
A10	0	0	CDX 1	0	N.C
A11	2	0	CDX 2	0	N.C
A13	2	2	CDX 3	0	N.C
B1	0	2	CDX 4	2	N.C
B2	0	0	Bicoid	2	0
B3	0	2			
B5	2	2	PAX	0	0
B6	3	3	PAX 6	0	0
B7	3	3	PAX 7	0	0
B8	1	2	Engrailed	1	1
B9	2	2	Goosecoid	0	0
C4	0	0			
C5	0	2	LIM a	3	2
C6	0	0	LIM b	3	2
C8	0	0	LH 1	0	0
C9	0	2	LH 2	0	0
C12	0	2			
C13	0	0	NK a	0	0
D3	0	2	NK b	0	0
D4	N.C	2	NK c	1	0
D8	2	2	NK d	1	1
D10	0	0	NK e	0	0
D11	0	0	ND f	0	0
			DLX a	0	0
			DLX b	2	0
			DLX c	0	0

Table 11 (continued)

Dispersed homeobox	Placenta	Amnion
EMX	0	0
OTX a	0	0
OTX b	0	0
EVE 1	0	0
EVE 2	0	0
GBX	1	0
MEI	2	0
MSX 1	0	0
MSX 2	3	0
PBX*	3	3
P HOX	1	0
PROX	0	1

Table 12

Serine Threonine phosphatase (STP) and Tyrosine phosphatase (TP) expression levels

Primer	Normal colon	HT-29	Caco-2	3T3
STP1 (740, 746)	1	2	1	1
STP2 (740, 747)	1	1	1	1
STP3 (740, 748)	0	1	0	0
STP4 (741, 746)	1	2	0	0
STP5 (741, 747)	2	2	2	1
STP6 (741, 748)	1	2	2	0
STP7 (742, 746)	1	2	2	2
STP8 (742, 747)	2	2	2	0
STP9 (742, 748)	2	2	2	2
STP10 (743, 749)	0	0	0	0
STP11 (743, 750)	0	0	0	0
STP12 (743, 751)	1	2	2	1
STP13 (744, 749)	0	0	0	0
STP14 (744, 750)	0	0	0	0
STP15 (744, 751)	0	0	0	0
STP16 (745, 749)	2	2	2	0
STP17 (745, 750)	2	1	2	0
STP18 (745, 751)	0	1	0	0
TP1 (705, 725)	0	0	0	0
TP2 (705, 726)	0	0	0	0
TP3 (705, 727)	0	0	1	0
TP4 (705, 728)	0	1	1	2
TP5 (706, 725)	0	0	0	0
TP6 (706, 726)	0	0	0	0
TP7 (706, 727)	1	2	2	2
TP8 (706, 728)	1	2	2	2
TP9 (708, 721)	2	2	2	2
TP10 (708, 722)	2	3	2	2
TP11 (708, 723)	2	3	2	2
TP12 (708, 724)	1	2	2	2
TP13 (709, 721)	2	2	2	2
TP14 (709, 722)	2	2	2	2
TP15 (709, 723)	2	2	2	2
TP16 (709, 724)	0	1	0	1
TP17 (710, 721)	2	3	2	1
TP18 (710, 722)	2	1	1	0
TP19 (710, 723)	1	2	1	1
TP20 (710, 724)	0	0	0	0

TP21 (711, 721)	2	1	2	2
TP22 (711, 722)	2	2	2	1
TP23 (711, 723)	2	2	2	1
TP24 (711, 724)	0	0	1	0

Table 13

Kinase expression levels

Primer	Normal colon	HT-29	Caco-2
STK31 (240, 251)	0	0	0
STK32 (240, 252)	2	0	0
STK33 (240, 253)	1	1	1
STK34 (241, 251)	2	0	1
STK35 (241, 252)	2	2	2
STK36 (241, 253)	2	1	1
STK37 (243, 251)	1	1	1
STK38 (243, 252)	1	1	1
STK39 (243, 253)	1	1	1
STK40 (245, 254)	2	1	2
STK41 (245, 255)	0	0	0
STK42 (246, 248)	1	1	0
STK43 (246, 249)	2	2	2
STK44 (246, 250)	1	1	0
STK45 (247, 248)	2	2	2
STK46 (247, 249)	2	2	2
STK47 (247, 250)	1	1	1
TK22 (401, 229)	0	0	0
TK23 (402, 229)	1	1	1
TK23 (403, 408)	2	2	2
TK24 (404, 229)	1	2	1
TK25 (405, 213)	1	2	1
TK26 (406, 409)	N.D.	1	1

Table 14

POU homeodomain expression levels

Primer	Normal colon	HT-29	Caco-2
POU1 (320, 326)	0	0	0
POU2 (320, 327)	1	1	1
POU3 (320, 328)	0	1	1
POU4 (321, 326)	0	0	0
POU5 (321, 327)	2	2	2
POU6 (321, 328)	2	2	2
POU7 (322, 326)	1	1	1
POU8 (322, 327)	1	2	2
POU9 (322, 328)	2	2	2
POU10 (323, 326)	1	1	1
POU11 (323, 327)	0	1	1
POU12 (323, 328)	1	2	1
POU13 (324, 326)	0	1	1
POU14 (324, 327)	2	1	2
POU15 (324, 328)	1	2	0

CLAIMS:

1. A method for determining the genetic proximity of a first cell and a second cell comprising the steps of:

(a) obtaining said first cell and said second cell;

(b) determining in said first cell and said second cell the pattern of expression of genes in a selected gene family;

(c) calculating a proximity index I, wherein

$$I = \frac{\sum_i \min(a_i, b_i)}{\sum_i \max(a_i, b_i)}$$

wherein a_i and b_i are the expression level of the gene i in the first and second cell, respectively, and the summations are performed over all genes i in the gene family,

a large proximity index indicating that said first cell and said second cell are genetically proximal to each other, a small proximity index indicating that said first cell and said second cell are genetically distant from each other.

2. A method according to Claim 1, wherein said pattern of expression of genes in said selected gene family is determined in at least one of said first cell and said second cell by the method of reverse transcriptase polymerase chain reaction.

3. A method according to Claim 2, wherein said reverse transcriptase polymer chain reaction utilizes one or more conserved primers.

4. A method according to Claim 3, wherein one or more of said one or more conserved primers is a primer selected from the list comprising Lists 1 to 19.

5. A method according to Claim 1 to 4, wherein said selected gene family is a set of homeobox genes.

6. A method according to Claim 5, wherein said homeobox genes are HOX genes.

7. A method according to any one of Claims 5 or 6, wherein said homeobox genes are dispersed homeobox genes.

8. A method according to Claim 1 to 4, wherein said selected gene family is a set of kinase genes.

9. A method according to any one of Claims 1 to 4, wherein said gene family is a set of protein phosphatase genes.

10. A method according to any one of Claims 1 to 4, wherein said gene family is a set of P450 enzyme genes.

11. A method according to any one of Claims 1 to 4, wherein said gene family is a set of steroid receptor superfamily genes.

12. A method according to any one of Claims 1 to 4, wherein said gene family is a set of cadhedrin superfamily genes.

13. A method according to any one of the preceding claims, wherein at least one of said first cell and said second cell is an amniotic cell.

14. Use of the method according to any one of Claims 1 to 13, for determining the effect of a selected treatment on a test cell wherein:

(a) said first cell is said test cell before having been subjected to said treatment;

(b) said second cell is said test cell after having been subjected to said treatment;

(c) and wherein said first cell and said second cell being genetically proximal to each other indicating that said treatment has no substantial effect on said test cell, and said first cell and said second cell being genetically distant from each other indicating that said treatment has a substantial effect on said test cell.

15. Use of the method according to any one of Claims 1 to 13, for determining whether a given test cell is transformed wherein:

(a) said first cell is said test cell;

(b) said second cell is an untransformed cell of the same cell type as said test cell;

(c) and wherein said first cell and said second cell being genetically proximal to each other indicating that said test cell is substantially untransformed, and said first cell and said second cell being genetically distant from each other indicating that said test cell is substantially transformed.

16. Use of the method according to any one of Claims 1 to 13, for detecting a selected genetic defect in a first individual wherein

(a) said first cell is a cell obtained from said first individual;

(b) said second cell is a cell of the same type as said first cell and obtained from a second individual not having the genetic defect;

(c) and wherein said first cell and said second cell being genetically proximal to each other indicating that said first individual substantially does not carry said genetic defect, and said first cell and said second cell being genetically distant from each other indicating that said first individual substantially carries said genetic defect.

17. Use of the method according to any one of Claims 1 to 13, for detecting a selected genetic defect in a first fetus wherein:

(a) said first cell is an amniotic cell obtained from said first fetus;

(b) said second cell is an amniotic cell obtained from a second fetus not having said genetic defect;

(c) and wherein said first cell and said second cell being genetically proximal to each other indicating that said first fetus substantially does not carry said genetic defect, and said first cell and said second cell being genetically distant from each other indicating that said first fetus substantially carries said genetic defect.

18. The use according to Claim 16 or 17, wherein said genetic defect is trisomy 21.

19. A kit for carrying out the method of any one of Claims 1-13, comprising:

(a) means for detecting the genes of said selected gene family expressed in said first cell and in said second cell;

(b) instructions for carrying out the method.

20. A kit for carrying out the method of any one of Claims 1 to 13 comprising:

(a) means for detecting the genes of said selected gene family expressed in said first cell;

(b) a catalogue providing the genes of said selected gene family expressed in said second cell;

(c) instructions for carrying out the method.

21. A kit according to Claim 19 for carrying out the method of any one of the preceding claims comprising:

(a) means for obtaining the mRNA from said first cell and said second cell;

(b) means for performing the reverse transcriptase polymer chain reaction on the mRNA obtained from said first cell and said second cell;

(c) means for detecting the genes of said selected gene family expressed in said first cell and in said second cell;

(d) instructions for carrying out the method.

22. A kit according to Claim 20 for carrying out the method of any one of Claims 1 to 13 comprising:

(a) means for obtaining the mRNA from said first cell;

(b) means for performing the reverse transcriptase polymer chain reaction on the mRNA obtained from said first cell;

(c) means for detecting the genes of said selected gene family expressed in said first cell;

(d) a catalogue providing the genes of said selected gene family expressed in said second cell;

(e) instructions for carrying out the method.

23. A method for obtaining cells capable of expressing an HB related desired property comprising the steps of:

(a) identifying a specific pattern of expression of HB genes in cells having a desired property;

(b) transfecting said identified set of HB genes into target cells lacking said desired property under conditions enabling expression of said HB cells in said target cells, said transfection resulting in expression of said desired property in said transfected cells.

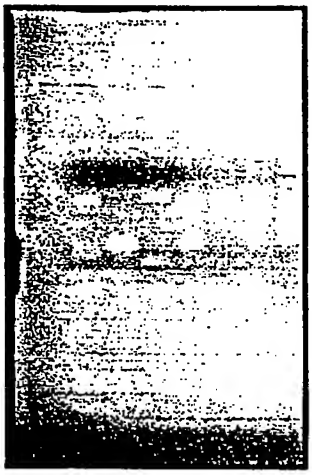

FIG.1A

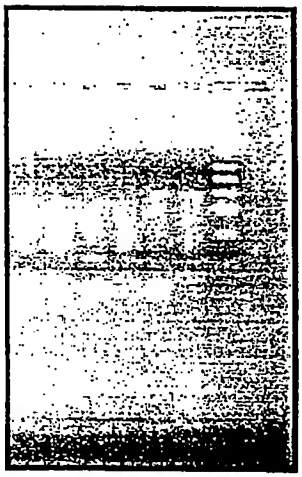

FIG.1B

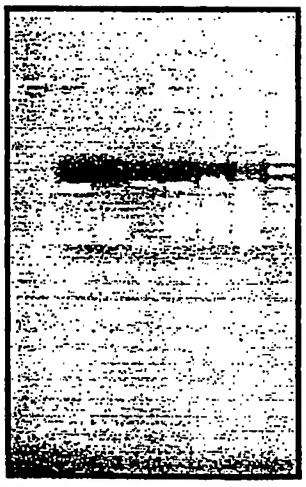

FIG.1C

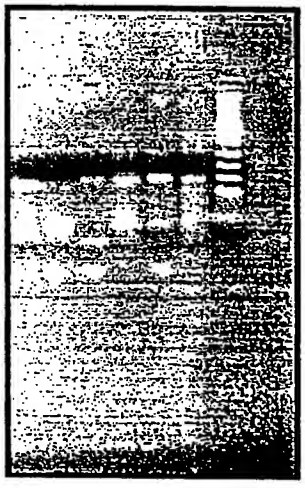

FIG.1D

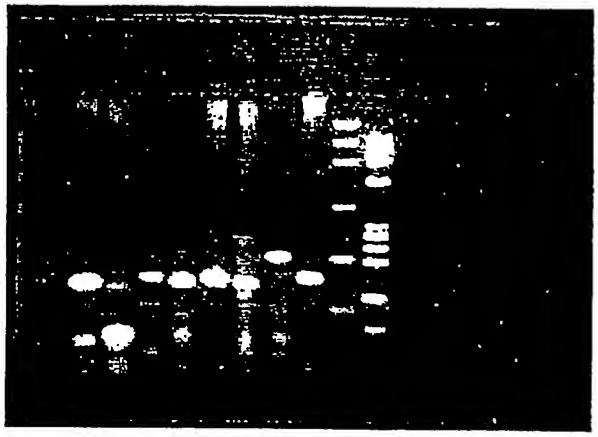

FIG.1E

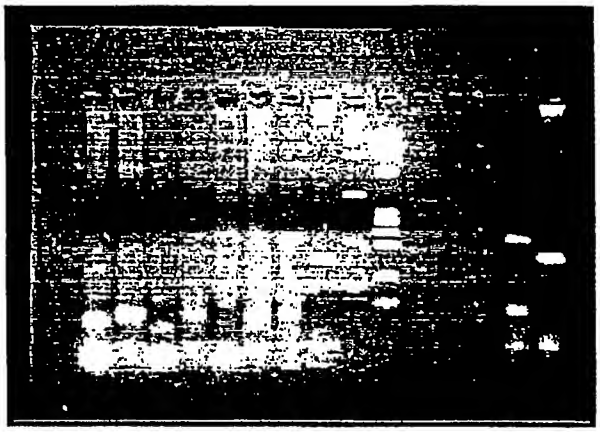

FIG. 1F

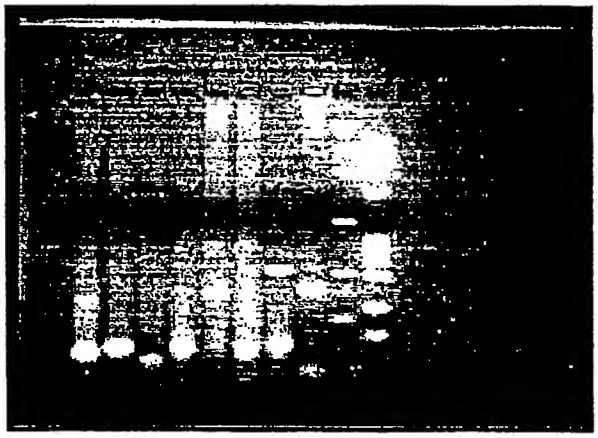

FIG. 1G

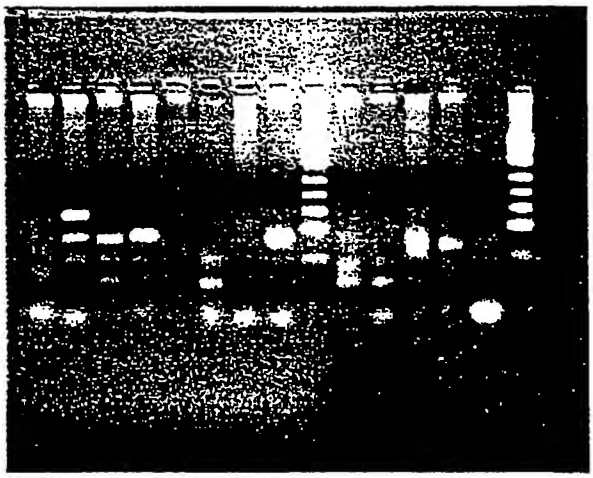

FIG.2A

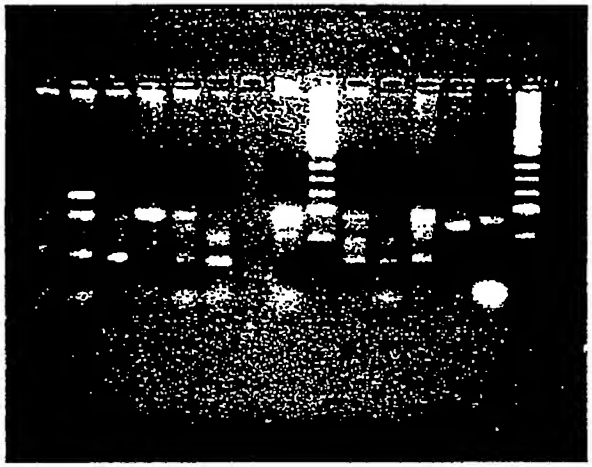

FIG.2B

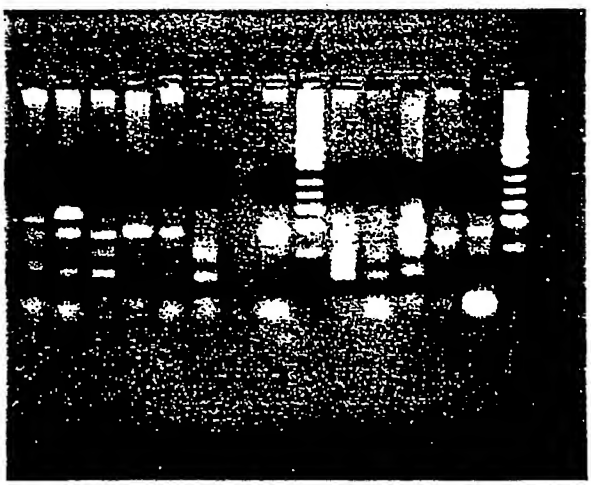

FIG.2C

FIG.3A

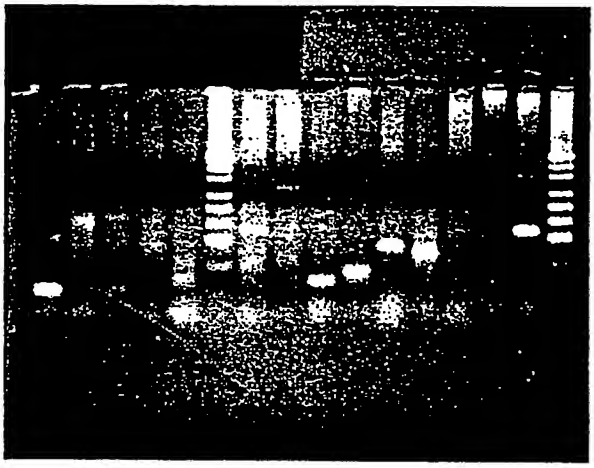

FIG.3B

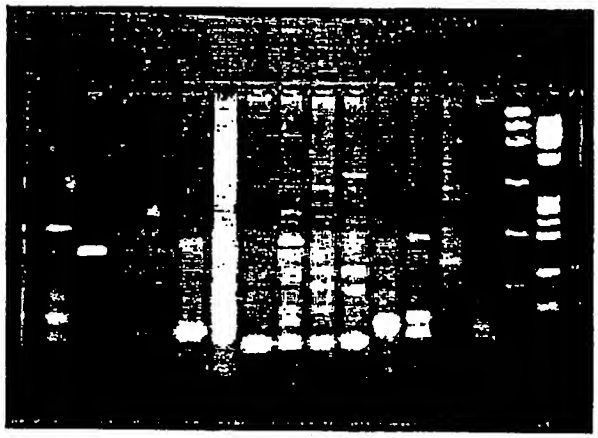

FIG.3C

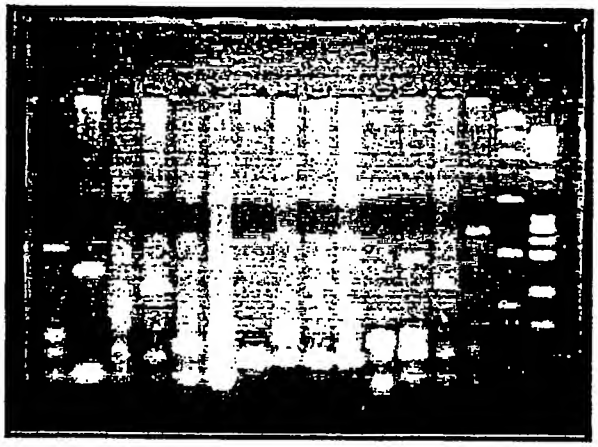

FIG.3D

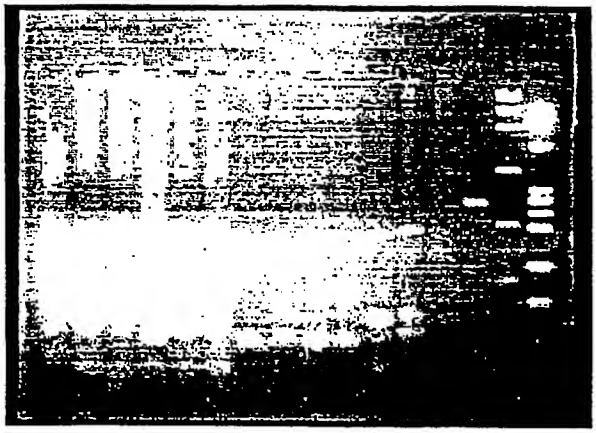

FIG.3E

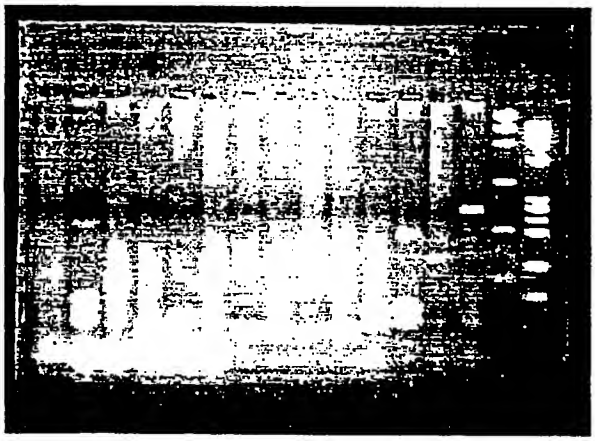

FIG.3F

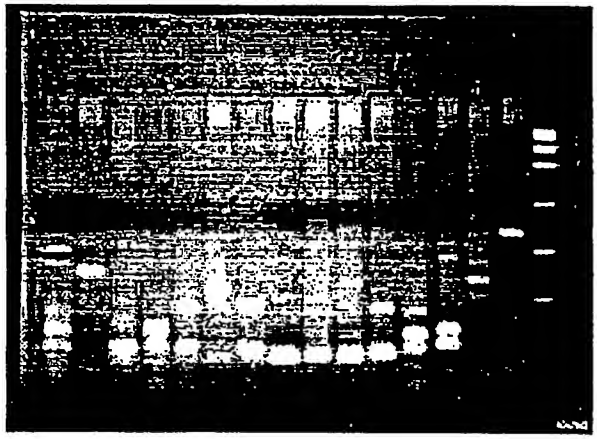

FIG.3G

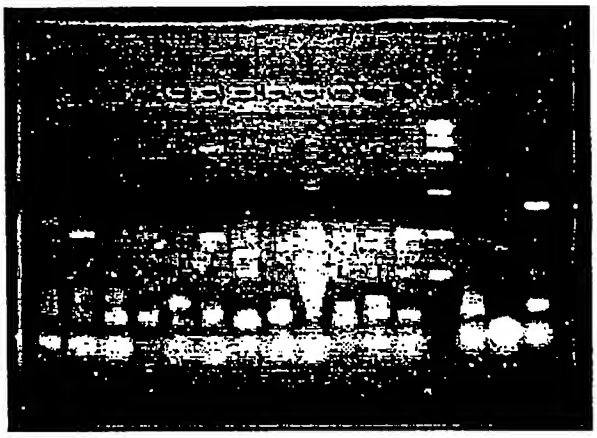

FIG.4A

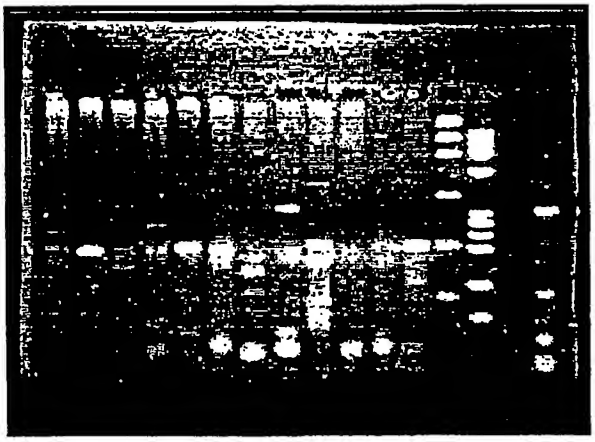

FIG.4B

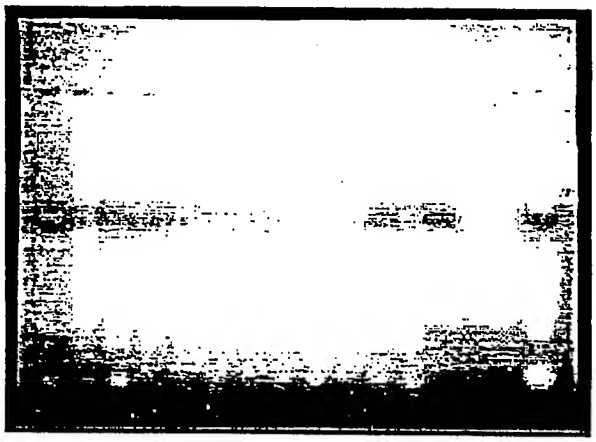

FIG.4C

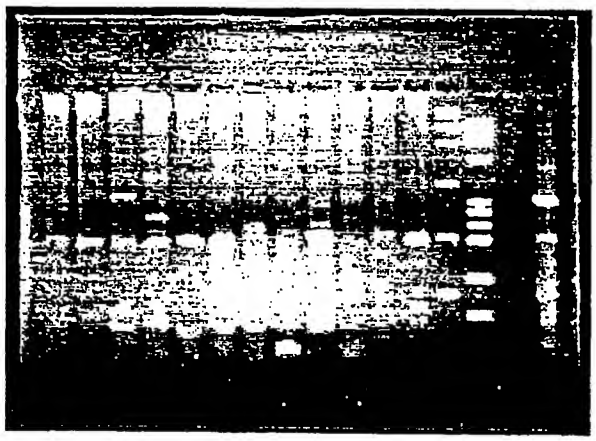

FIG.4D

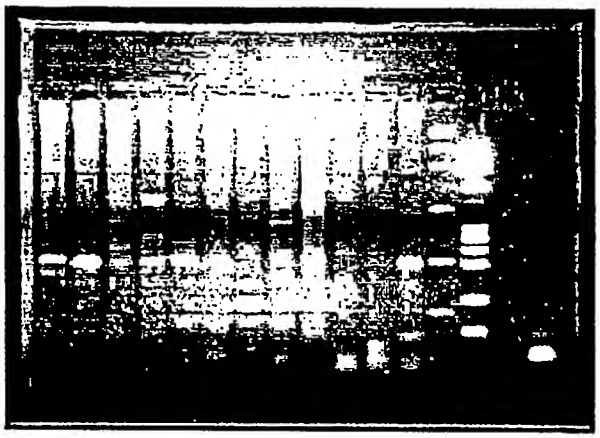

FIG.4E

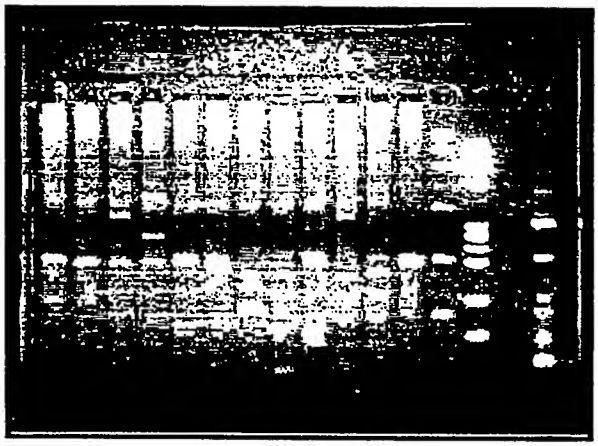

FIG.4F

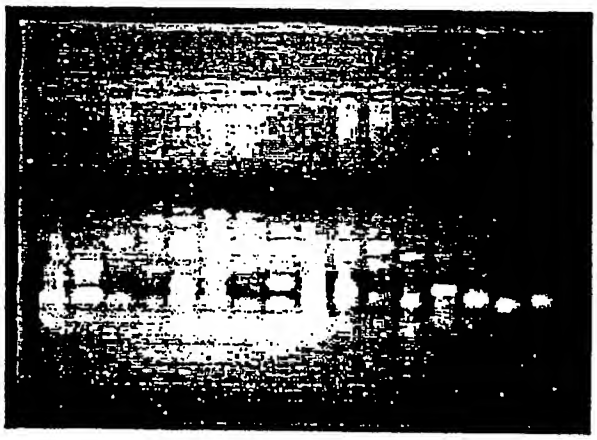

FIG.5A

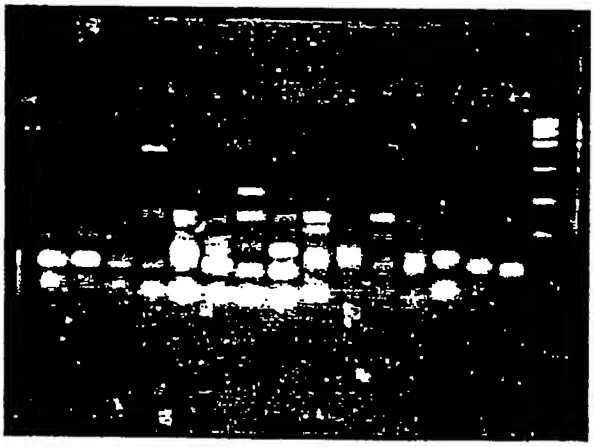

FIG.5B

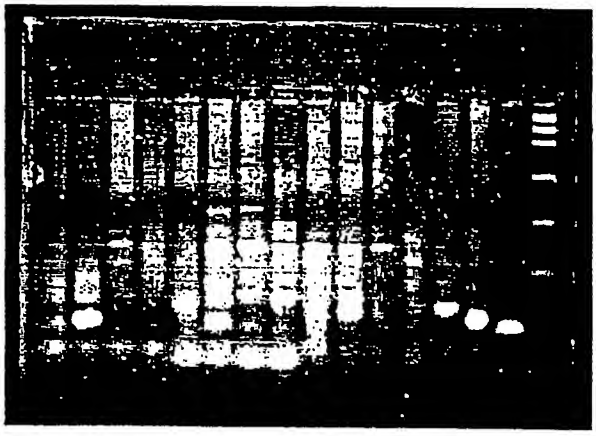

FIG.5C

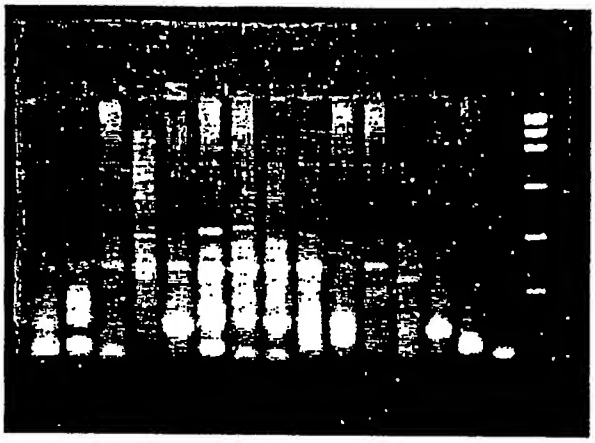

FIG.5D

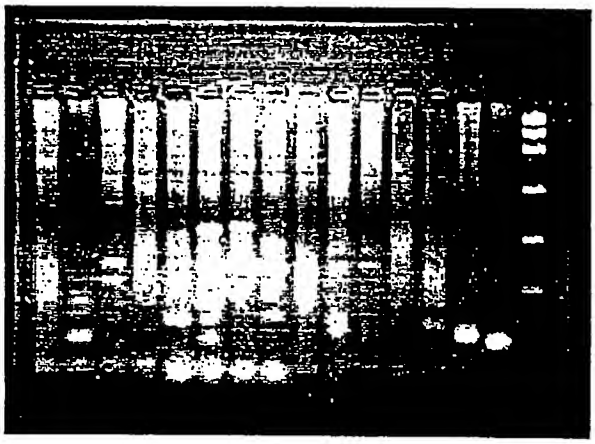

FIG.5E

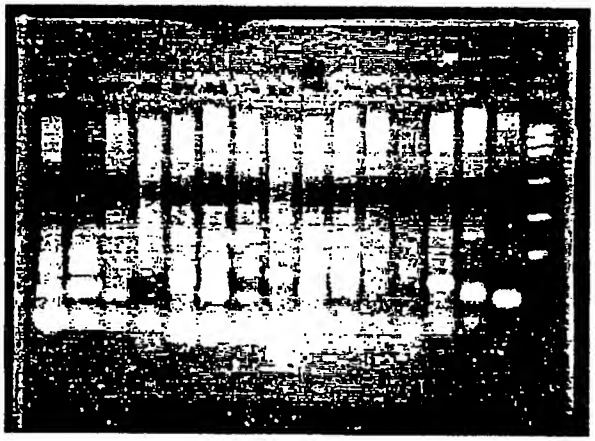

FIG.5F

FIG.5G

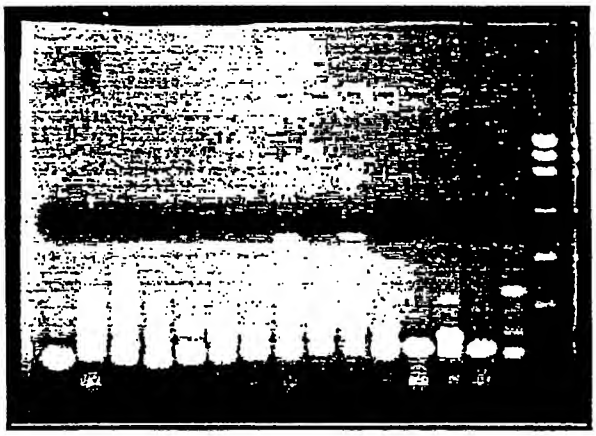

FIG.6

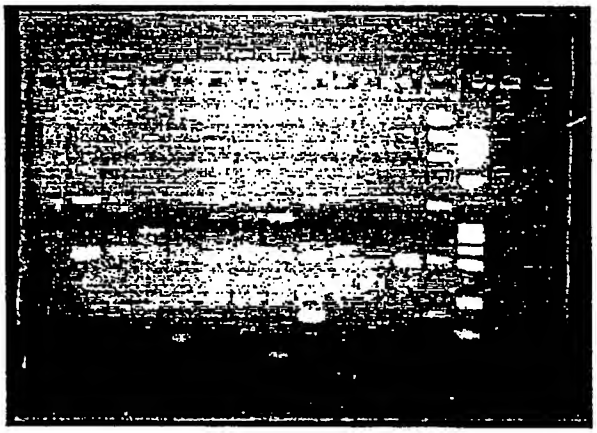

FIG.7A

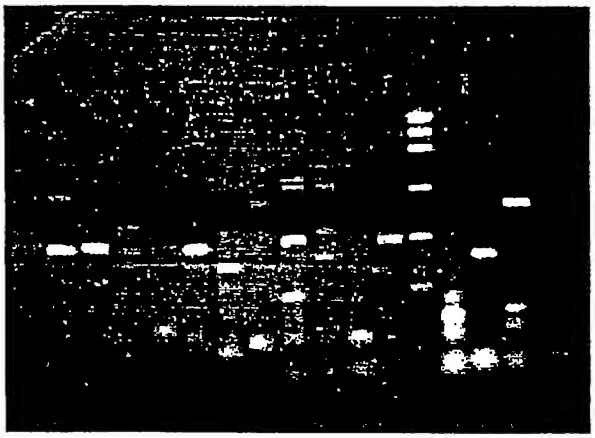

FIG.7B

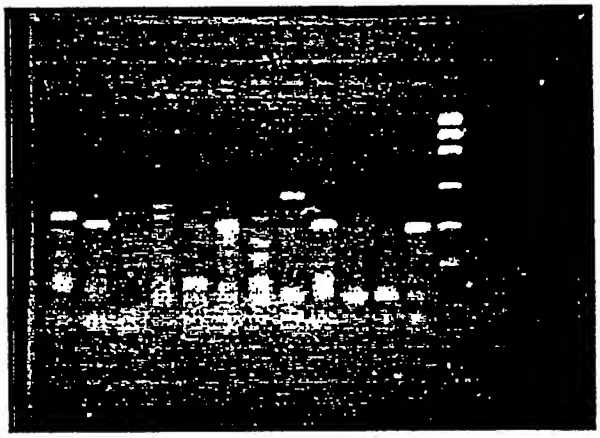

FIG.7C

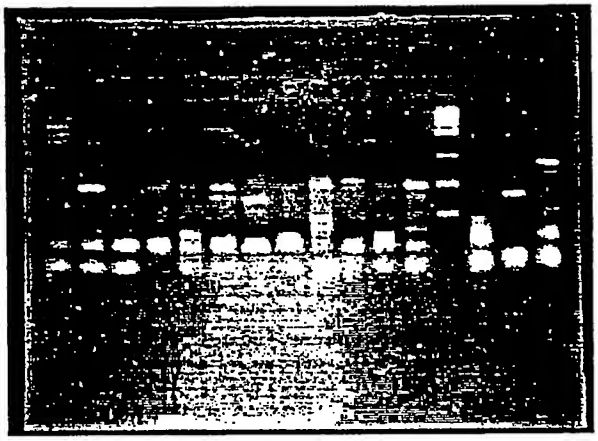

FIG.7D

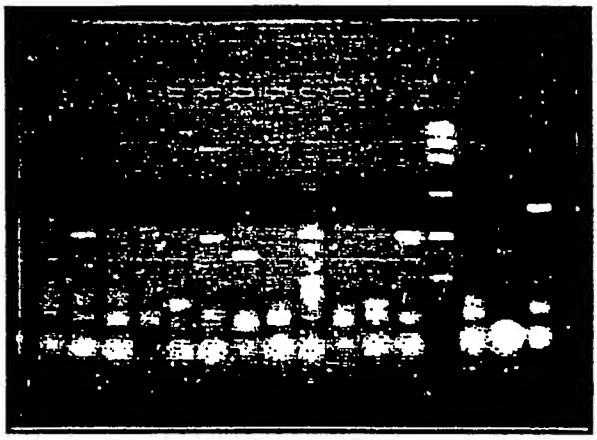

FIG.7E

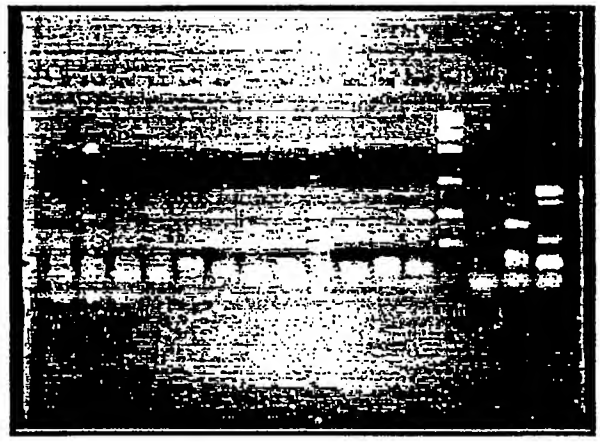

FIG.7F

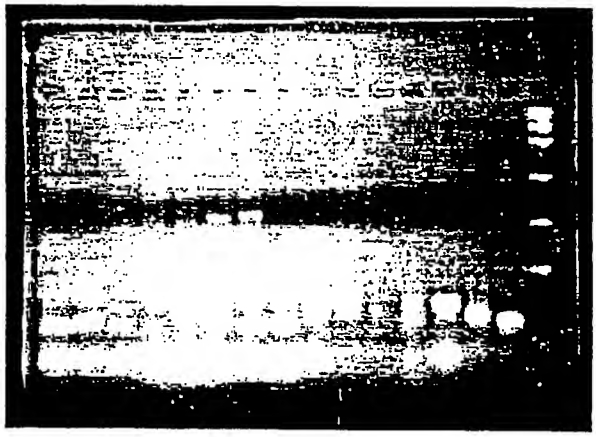

FIG.8A

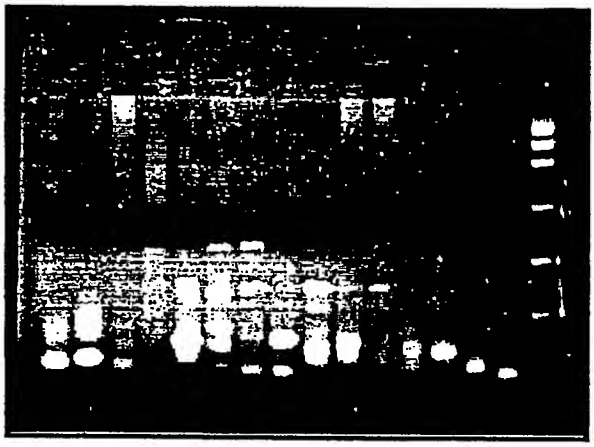

FIG.8B

FIG.8C

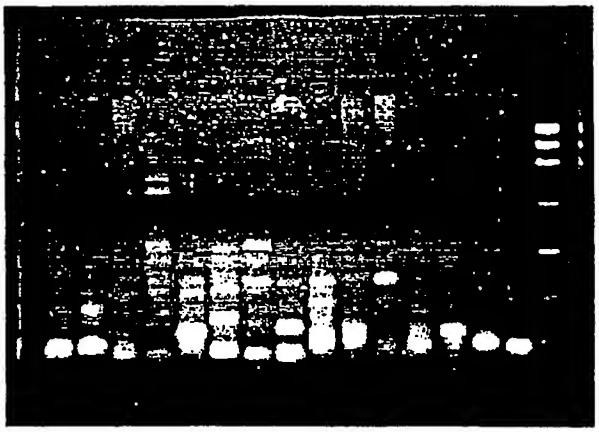

FIG.8D

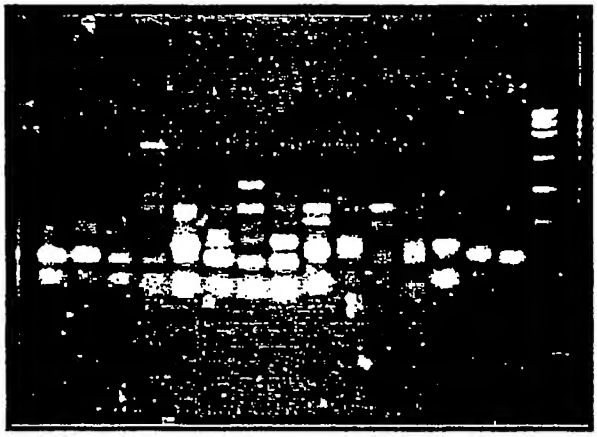

FIG.8E

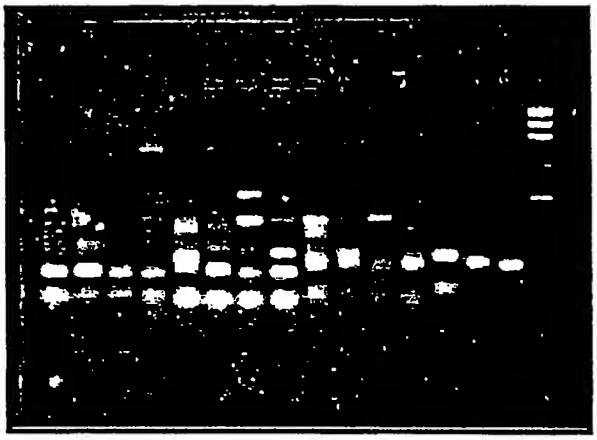

FIG.8F

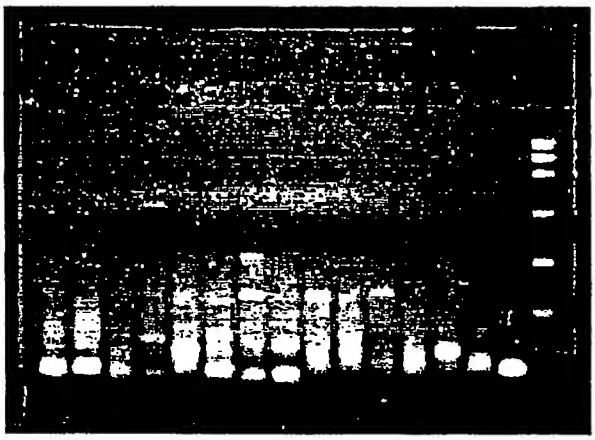

FIG.8G

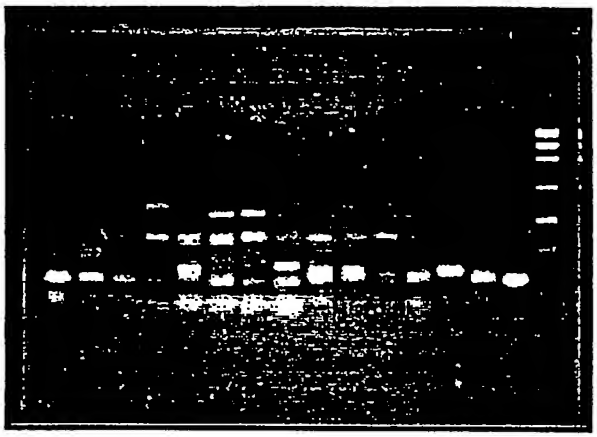

FIG.8H

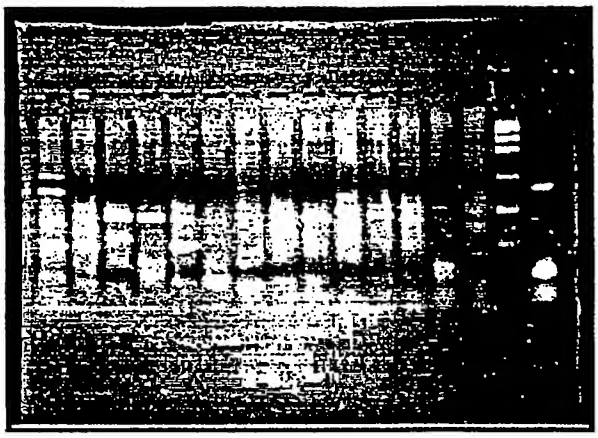

FIG.9A

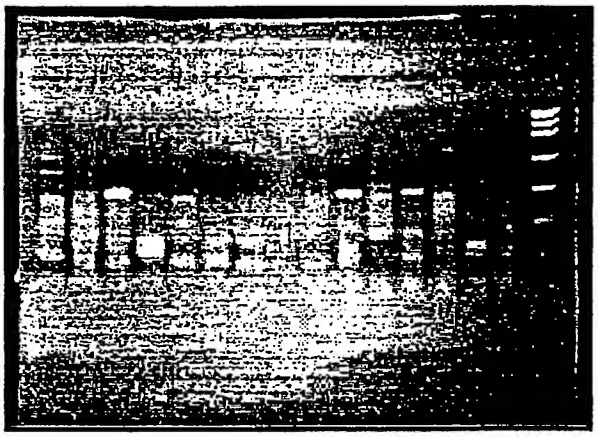

FIG.9B

FIG.9C

INTERNATIONAL SEARCH REPORT

Inte. .ional Application No
PCT/IL 98/00625

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12Q1/68

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 96 34985 A (SUGEN INC) 7 November 1996 see the whole document ---	1-22
A	EP 0 743 367 A (MAX PLANCK GESELLSCHAFT) 20 November 1996 see the whole document ---	1-22
A	WILKIE T M ET AL: "CLONING MULTIGENE FAMILIES WITH DEGENERATE PCR PRIMERS" METHODS: A COMPANION TO METHODS IN ENZYMOLOGY, vol. 2, no. 1, 1 February 1991, pages 32-41, XP000565446 see the whole document --- -/--	1-22

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
8 June 1999	21/06/1999

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Osborne, H

1

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

Inte .ional Application No
PCT/IL 98/00625

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
A	VIDER B ET AL : "Human colorectal carcinogenesis is associated with deregulation of homeobox gene expression" BIOCMEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS, vol. 232, no. 1, March 1997, pages 742-48, XP002104685 cited in the application see the whole document	1-23
A	WO 95 20681 A (INCYTE PHARMA INC) 3 August 1995 see the whole document	1-22
A	WO 95 21944 A (SMITHKLINE BEECHAM CORP ;ROSENBERG MARTIN (US); DEBOUCK CHRISTINE) 17 August 1995 see the whole document	1-22

1

INTERNATIONAL SEARCH REPORT

Information on patent family members

Int. ational Application No
PCT/IL 98/00625

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9634985 A	07-11-1996	US 5830648 A	03-11-1998
		AU 5636996 A	21-11-1996
		CA 2220300 A	07-11-1996
		EP 0824597 A	25-02-1998
EP 0743367 A	20-11-1996	DE 19518505 A	21-11-1996
		JP 8308598 A	26-11-1996
		US 5876932 A	02-03-1999
WO 9520681 A	03-08-1995	US 5840484 A	24-11-1998
		AU 688465 B	12-03-1998
		AU 1694695 A	15-08-1995
		BG 100751 A	31-07-1997
		BR 9506657 A	16-09-1997
		CA 2182217 A	03-08-1995
		CN 1145098 A	12-03-1997
		CZ 9602189 A	14-05-1997
		EP 0748390 A	18-12-1996
		FI 962987 A	26-09-1996
		JP 9503921 T	22-04-1997
		LV 11696 B	20-08-1997
		NO 963151 A	27-09-1996
		PL 315687 A	25-11-1996
		HU 75550 A	28-05-1997
WO 9521944 A	17-08-1995	EP 0743989 A	27-11-1996
		JP 9508800 T	09-09-1997